

Supplementary Information

QUANTIFYING UNDERESTIMATES OF LONG-TERM UPPER-OCEAN WARMING

Paul J. Durack¹, Peter J. Gleckler¹, Felix W. Landerer² and Karl E. Taylor¹

¹Program for Climate Model Diagnosis and Intercomparison, Lawrence Livermore National
Laboratory, Livermore, California, USA

²Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA

Draft: Monday, 14 July 14

Subsections:

1. Published estimates of observed ocean heat content (OHC) change
2. Global ocean ventilation – observations and models
3. Total steric change and sea surface height
4. Signal and noise across CMIP5 experiments – unforced internal variability
5. Full CMIP ensemble supporting analysis
6. Data sources
7. Data preparation and analysis
8. Supplementary tables
9. Supplementary references

Supplementary Figures: S1, S2a, S2b, S3, S4, S5, S6, S7, S8, S9, S10 and S11

Supplementary Tables: S1 and S2

1. Published estimates of observed Ocean Heat Content (OHC) change

There are many analyses that have examined and discussed changes to long-term ocean heat content over the observed period (mid-20th century to the present). Such studies include contributions from AchutaRao *et al.* (2006, 2007), Barnett *et al.* (2001, 2005), Church *et al.* (2004, 2011), Domingues *et al.* (2008), Ishii *et al.* (2003, 2006), Ishii & Kimoto (2009), Gleckler *et al.* (2012), Good *et al.* (2013), Gouretski *et al.* (2012), Harrison & Carson (2007), Levitus *et al.* (2000, 2005, 2009, 2012), Lyman & Johnson (2008, 2014), Marcos & Amores (2014), Palmer *et al.* (2007, 2009), Pierce *et al.* (2006, 2012) and Smith & Murphy (2007) amongst others. A number of other studies have generated analyses over the more recent period (Boening *et al.*, 2012; Lyman *et al.*, 2010; Johnson *et al.*, 2013; Willis *et al.*, 2004; von Schuckmann & Le Traon, 2011), however we do not discuss these here.

Due to the spatial and temporal sparsity of the observed data, these studies have considered two key approaches to generating global estimates:

Spatially incomplete “representative global average” analysis

A subsampled (or spatially incomplete) analysis attempts to generate a “representative global average” by using existing observations and extrapolating an area-weighted sample average over unsampled regions. Importantly this technique leads to a global ocean sampled volume that varies with time, and a key assumption is that the ocean heat content (OHC) in unsampled regions (most of the Southern Hemisphere [SH]) is changing at the same rate as the sampled regions. The Palmer *et al.* (2007) study is an example of a “representative global average” analysis, and we present some example results over four years (1970, 1980, 1990, 2000) for which a “representative global average” infilling was undertaken (Figure S1). It is clear that for the four years plotted, almost all “representative global average” information is obtained from the North Atlantic and northwestern Pacific regions and this data is used to infill most of the SH, where only few observations exist. This implies that the “representative global average” approach is not appropriate to investigate OHC change partitioning between hemispheres.

Spatially complete “infilled” analysis

Analyses that follow a spatially complete approach use various methods to infill regions where available observations do not provide coverage. One of the more common methods employed is objective analysis (e.g. Levitus, 1982, 1984; Ishii *et al.*, 2003, 2005) which constructs either climatological annual, seasonal or monthly mean gridded fields. A disadvantage of these analyses is that where observations do not provide coverage, “infilled” climatological values or zero anomaly values are used. A number of studies have noted that such “infilled” analyses tend to underestimate changes in poorly sampled regions (Gille, 2002, 2008; Gregory *et al.*, 2004; Gouretski & Koltermann, 2007) or may underestimate inherent climate variability (Harrison & Carson, 2007).

67 More recent estimates (Domingues *et al.*, 2008; Lyman & Johnson, 2008, 2014) have
68 attempted to address the poor spatio-temporal coverage by relying on the coincident
69 sea surface height (SSH) estimates along with the Argo array, that both provide near
70 global coverage over the more modern period. In the case of Durack & Wijffels (2010)
71 these estimates depend on the modern Argo period to reduce seasonal and spatial
72 sampling biases in the historical archive, and make the assumption that the Argo period
73 is a valid representation of the global ocean's spatial structure and seasonal cycle for the
74 historical period.

75 As a consequence of these optimisations, the newer analyses tend to provide larger
76 magnitudes of change, particularly for the poorly sampled regions. The 4 “infilled”
77 analyses assessed in this study for which maps are available are presented in Figure S2a,
78 for reference multi-model mean (MMM) results from the CMIP3 and CMIP5
79 experiments are presented in Figure S2b.

80 A detailed overview of ocean temperature observations, with a key focus on expendable
81 bathythermograph (XBT) bias corrections is provided in Abraham *et al.* (2013). To date,
82 there have been a number of studies published on addressing fall-rate and temperature
83 bias corrections for XBT and Mechanical BathyThermograph (MBT) temperature
84 observing platforms, and include contributions from Wijffels *et al.* (2008), Gouretski &
85 Reseghetti (2010), Good (2011), Hamon *et al.* (2012), Gouretski (2012) and Cowley *et al.*
86 (2013).

87 A very recent study by Cheng & Zhu (2014) has highlighted a new source of uncertainty
88 relevant to observed estimates of OHC change. Their study considered historical
89 observations, and found that due to the poor vertical resolution of these data complex
90 vertical and geographical temperature biases are likely. They concluded that the 0 - 700
91 m OHC is likely biased high, however geographically this bias is complex, with positive
92 values between 30°S – 30°N and a cold bias at higher latitudes in both hemispheres,
93 which is sensitive to the shape (concave or convex) of the vertical temperature profile.

94 Such new insights, along with the hemispheric heat partitioning constraint which is the
95 focus of the current work, will need to be considered by future analyses which aim to
96 construct improved estimates of global OHC change.

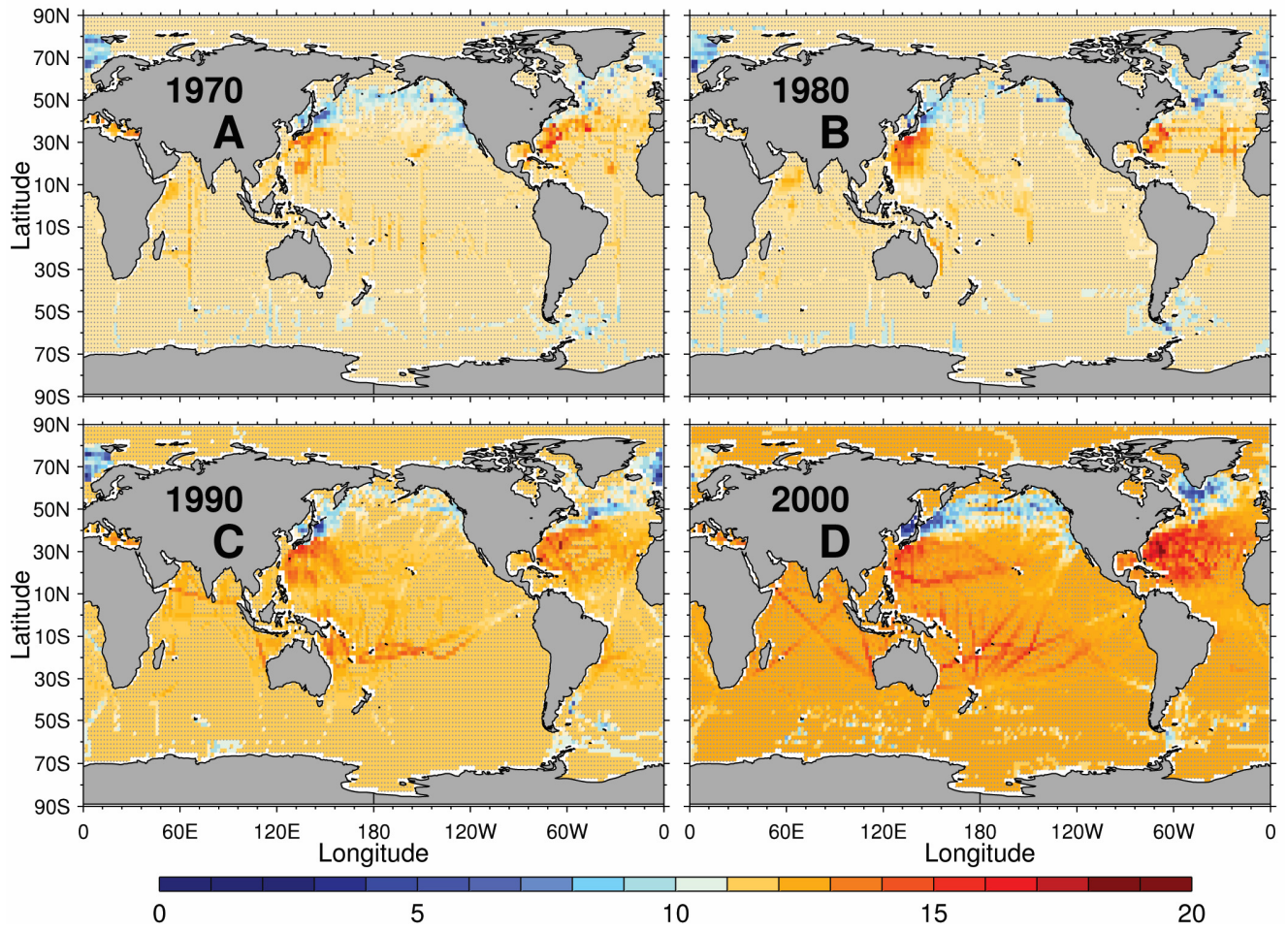


FIG. S1. Annual mean temperature maps (units are degrees Celsius) obtained across the 12 months for target years A) 1970, B) 1980, C) 1990 and D) 2000 using the “representative average” methodology of Palmer *et al.* (2007). For each year of available gridded data noted above, a representative annual mean is calculated by averaging all available monthly values to yield an annual mean value. From this incomplete annual mean field a “representative global average” is calculated using area-weights from the observational grid, and then for each unobserved grid point the “representative global average” is inserted. Stippling denotes regions where the “representative global average” has been used to replace a missing value in the original gridded data. We note that for each of the selected years, Northern Hemisphere (NH) values dominate the “representative global average” and that there is very little Southern Hemisphere (SH) representative information available.

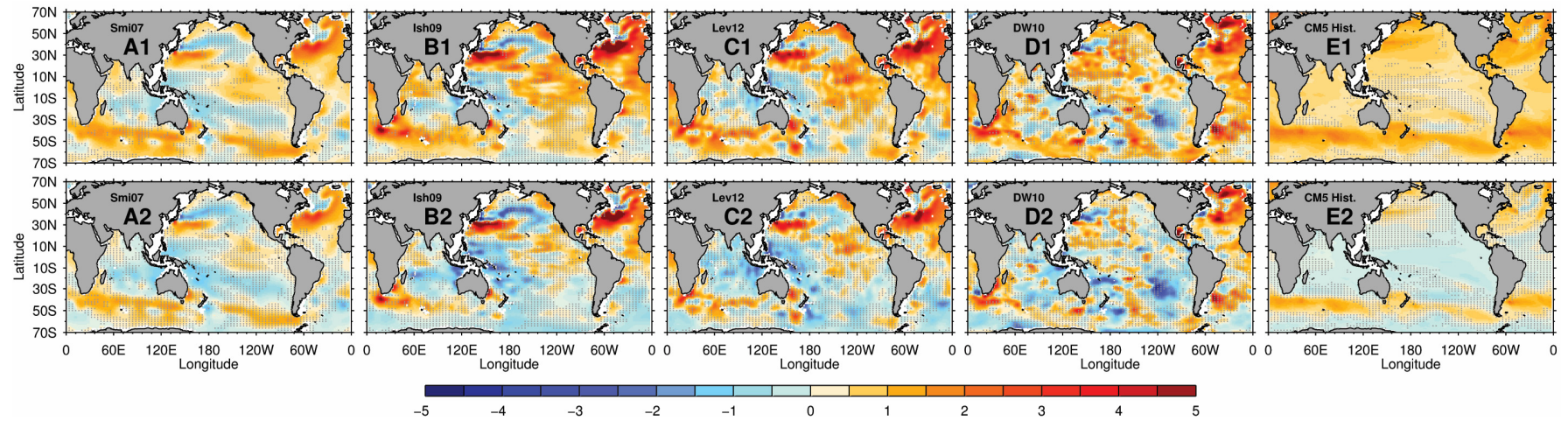


FIG. S2a. 35-year trends in ocean heat content for the upper 0-700 dbar with the global mean included (A1-E1; top row), and with the global mean removed (A2-E2; bottom row). Units are $\text{J} \times 10^3 \text{ kg}^{-1} 35\text{yrs}^{-1}$ ($4 \approx$ depth-averaged warming of $1^\circ\text{C} 35\text{yrs}^{-1}$). Observational maps show the results from Smith & Murphy (2007; 1970-2004, A), Ishii & Kimoto (2009; 1970-2004, B), Levitus *et al.* (2012; 1970-2004, C), Durack & Wijffels (2010; 1970-2008 – scaled to represent 35yrs^{-1} , D) and the CMIP5 historical multi-model mean (MMM; 1970-2004, E). Stippling is used to mark regions where the 4 observational estimates do not agree in their sign (A, B, C, D) and where less than 75% of CMIP5 models do not agree in sign with the averaged map obtained from the ensemble (E). See Figure S1b for each independent CMIP experiment MMM result.

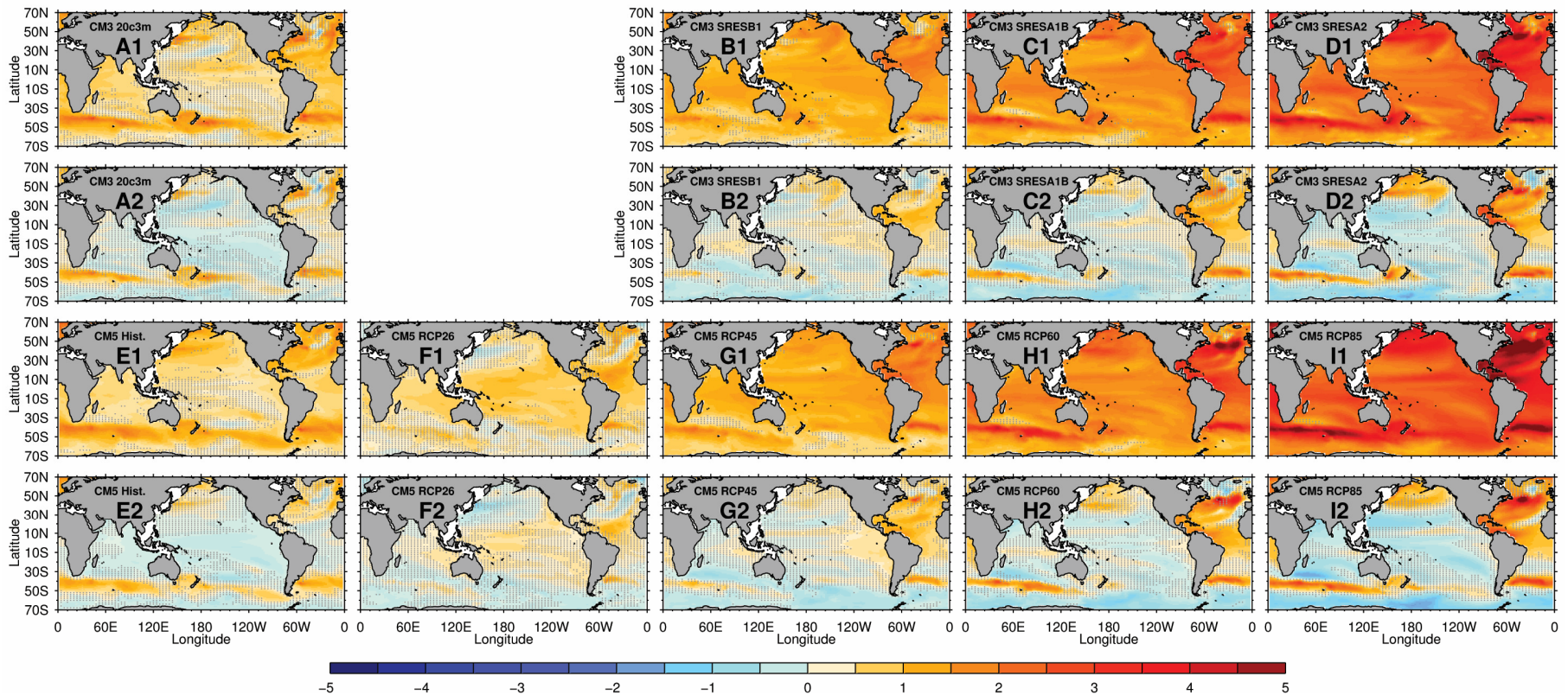


FIG. S2b. 35-year trends in ocean heat content for the upper 0-700 dbar with the global mean included (top and third rows), and with the global mean removed (second and bottom rows). Units are $\text{J} \times 10^3 \text{ kg}^{-1} 35\text{yrs}^{-1}$ ($4 \approx$ depth-averaged warming of $1^\circ\text{C} 35\text{yrs}^{-1}$). Multi-model mean (MMM) maps show the results from CMIP3 20c3m (A – 1970-1999 scaled to represent 35yrs^{-1}), CMIP3 SRESB1 (B – 2065-2099), CMIP3 SRESA1B (C – 2065-2099), CMIP3 SRESA2 (D – 2065-2099), CMIP5 historical (E – 1970-2004), CMIP5 RCP26 (F – 2065-2099), CMIP5 RCP45 (G – 2065-2099), CMIP5 RCP60 (H – 2065-2099) and CMIP5 RCP85 (I – 2065-2099). Following Figure S2a, panels denoted “1” show the absolute result (e.g. A1) whereas panels denoted “2” show maps with the area-weighted global mean removed (e.g. A2). Stippling is used to mark regions where less than 75% of CMIP models do not agree in sign with the averaged map obtained from each ensemble (A-I).

2. Global ocean ventilation – observations and models

The rate of observed ocean ventilation (transport of surface waters leaving the surface formation site moving into the ocean interior) can strongly affect estimates of long-term OHC change. However, long-term data coverage is insufficient to constrain OHC trends, let alone the highly temporally variable ocean ventilation. Additionally, it is likely that observed ocean ventilation rates are not static over the 1970 to 2004 period of analysis (Waugh *et al.*, 2013), and may be trending upwards due to wind changes in some regions (Abram *et al.*, 2014).

Ocean ventilation is a difficult process to quantify, as ventilation is a response to dynamical ocean processes affected by ocean stratification, mixed-layer dynamics, vertical diffusion, convection, eddy-flux and surface ocean fluxes along with responses to windstress forcing (e.g. Bryan *et al.*, 2006; Downes *et al.*, 2010; Capotondi *et al.*, 2012; Downes & Hogg, 2013; Morrison *et al.*, 2013; Salleé *et al.*, 2013).

In models, many of these processes are parameterized as they occur at sub-grid spatial scales. This simplification is necessary to allow models to simulate the complete global ocean. In practice, due to poor temporal and spatial observational coverage, oceanographers attempt to quantify ocean ventilation using anthropogenic transient tracers, for which a well-known time history is available. Chlorofluorocarbons (CFCs) are such a common tracer, and are particularly useful due to the chemical- and biological-inert nature of the CFC-11 and CFC-12 species (Bullister, 1989; Bullister & Tanhua, 2010). These tracers are well mixed in the atmosphere for both hemispheres (Figure S3). Spatial maps showing CFC-11 inventories for the ocean (Figure S3), in response to ocean ventilation, qualitatively agree with the key regions of ocean heat uptake (Southern Ocean and North Atlantic) in the Southern Ocean and North Atlantic as captured in models (Figure S2b) and the sparse observations (Figure S2a).

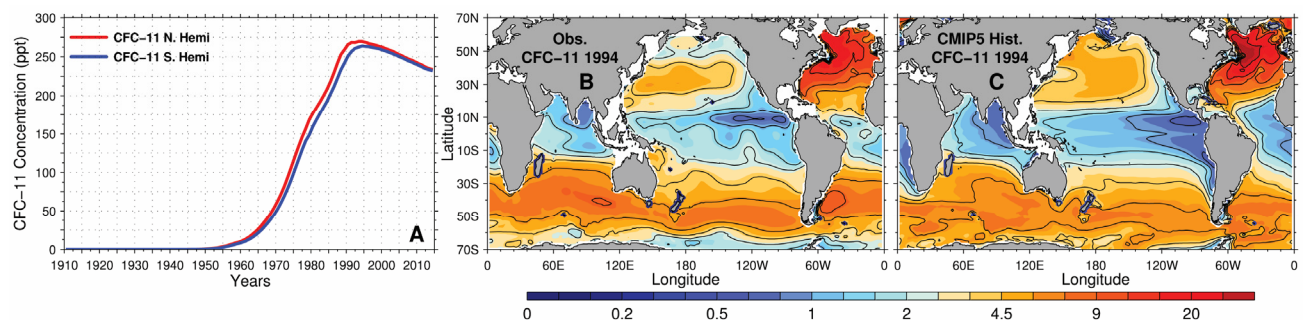


FIG. S3. The historical atmospheric timeseries and time-integrated oceanic concentrations of chlorofluorocarbon-11 (CFC-11; Bullister, 2014). (A) CFC-11 atmospheric mole fraction in both hemispheres. Mapped vertically integrated ocean CFC-11 concentrations for (B) observations (0-depth of deepest samples) and (C) for a representative ensemble mean from CMIP5 (4 models; 0-full depth). Concentrations are shown for the representative year 1994 (the median year of observational data) and the 1994 annual mean (from monthly data) concentrations from available CMIP5 models. Units are moles/km², highest values are red, lowest are blue.

A number of studies have compared inferred ocean ventilation rates between models and observations (e.g. Dixon *et al.*, 1996; Dutay *et al.*, 2002; Danabasoglu *et al.*, 2009; Shao *et al.*, 2013). These studies suggest that a qualitative agreement exists between modelled and observed CFC-11 uptake for the global ocean, however there is a large range ($\pm 30\%$) in the resolved global inventory in models, which was attributed to differences in high latitude ocean ventilation (Dutay *et al.*, 2002). The Dutay *et al.* (2002) study focused on the CMIP2-generation (Meehl *et al.*, 2000) models, which contributed to the Ocean Carbon-cycle Model Intercomparison Project (OCMIP). This study found the largest discrepancies between modelled and observed CFC concentrations, were located in the high latitudes and were sensitive to subgrid-scale parameterizations, along with isopycnal diffusion and eddy-induced velocity parameterizations.

When comparing the CMIP2 and CMIP5 model suites, large improvements to simulation realism have been achieved (e.g. Flato *et al.*, 2013) by addressing the sensitivities noted above, and increasing horizontal and vertical resolution. It is anticipated that such model improvements would lead to a more realistic ocean ventilation simulation when compared to observed estimates, and as seen in Figure S3. The sub-suite of the CMIP5 models which provide CFC-11 concentration data (CESM1-CAM5, GFDL-CM3, GFDL-ESM2G, NorESM1-ME) replicate observed spatial patterns well (contrast Figure S3B versus C). The Southern Hemisphere (SH) fractional contribution to the global upper OHC expressed by this subset range from 0.46 to 0.63, with a multi-model mean (MMM) of 0.57 and so these four models provide a very representative subset of the full suite shown in Figure 4 (MMM = 0.59). This result suggests that the ventilation locations and rates in CMIP5 Historical models are realistic.

3. Total steric change and sea surface height

Observed sea surface height (SSH) from highly accurate satellite-based altimeters have been used to inform spatial infilling of ocean heat content anomaly (OHCA) and thermosteric sea-level estimates in previous works (Church *et al.*, 2004; Domingues *et al.*, 2008) or to provide insights into uncertainty estimates (Lyman & Johnson, 2008, 2014). These studies leverage off the strong correspondence between full-depth steric changes and the resultant SSH, as well as the near global, high spatio-temporal resolution of continuous SSH measurements since late 1992.

To validate our analysis and corroborate our approach of considering numerous observed and modelled quantities, we investigated the correspondence between large-scale integrated estimates of OHC change, and both thermosteric and total (thermosteric and halosteric) steric ocean responses. Due to the limitation of SSH observations (not available before October 1992), these comparisons were undertaken only within the CMIP5 model suite over the 35-year period of assessment and are shown in Figure S4.

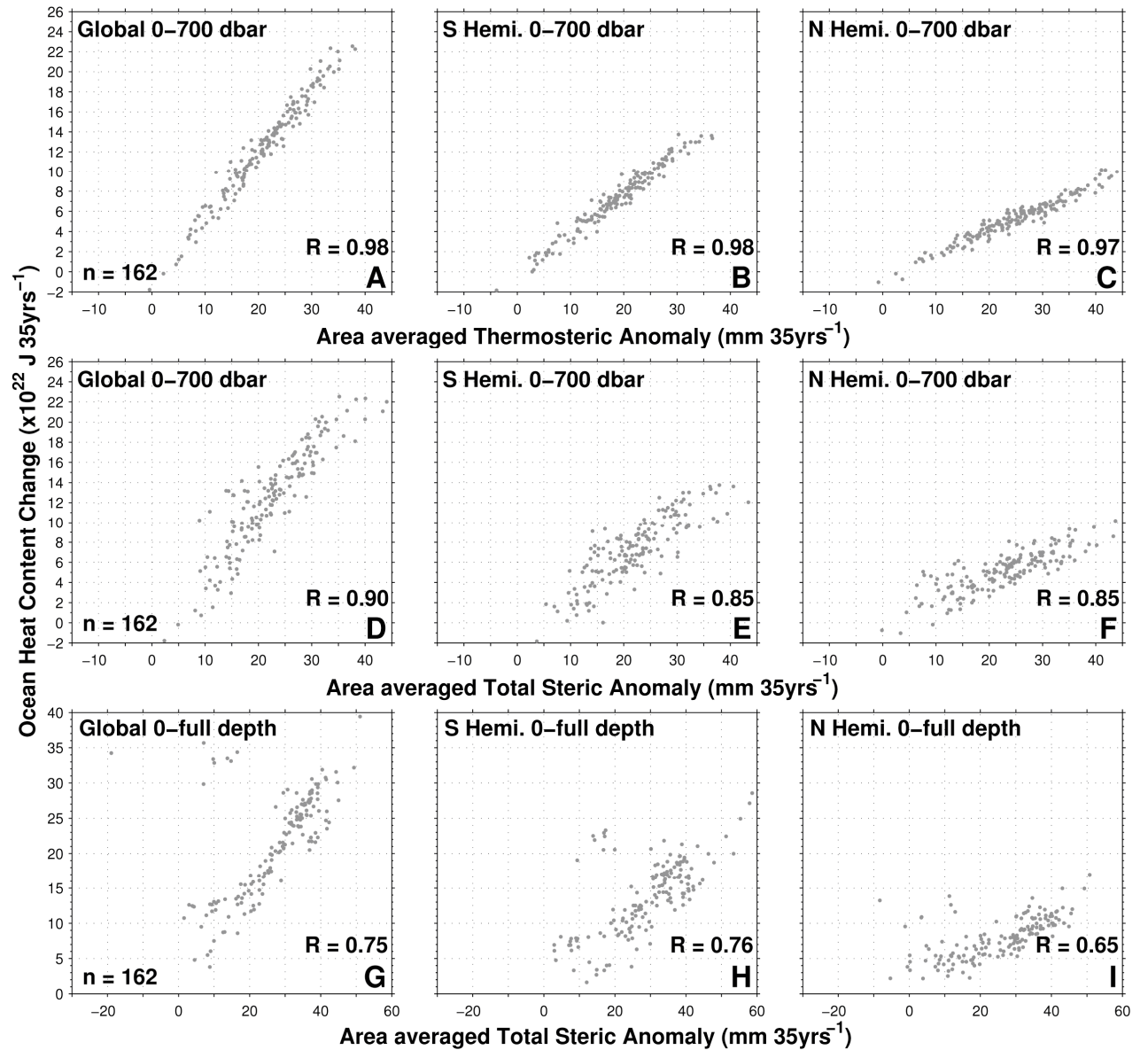


FIG. S4. Comparison of integrated quantities obtained from 35-year trends of ocean heat content (OHC) changes contrasted with (A, B, C) 0-700 dbar thermosteric anomalies, (D, E, F) 0-700 dbar total (thermosteric and halosteric) steric anomalies and (G, H, I) 0-full depth total steric anomalies in available CMIP5 Historical simulations. We note the 0-full depth quantities are not drift-corrected. When drift-correction is undertaken correlations improve to 0.98, 0.95 and 0.83 for panels G, H and I respectively (see Figure 2D, E, F). 9 simulations from the GISS model suite were removed from the analysis presented in panels G, H, I due to spuriously large deep ocean heating (global integral $> 10^{23}$ J); these issues are normally resolved by drift-correction as undertaken in Figure 2 (panels D-F).

Following previous studies (Church *et al.*, 2004; Domingues *et al.*, 2008; Lyman & Johnson, 2008, 2014), we use near-global SSH observations alongside modelled estimates to investigate consistency on hemispheric-scales between observed and

modelled long-term OHC. Modelled estimates of global mean sea-level change over the historical period have been shown to be equivalent to observed estimates when considerations for excluded land-ice contributions are taken into account (Church *et al.*, 2013). The spatial correspondence between these quantities is captured well in the CMIP5 models (also shown for integrated global and hemispheric totals in Figure S4), and we show maps of the various simulated vertical components of OHC (0-700 dbar and 0-full depth) along with the total steric anomaly (calculated from temperature and salinity fields) and modelled SSH (CMIP5 variable “ZOS”, sea surface height above geoid) in Figure S5.

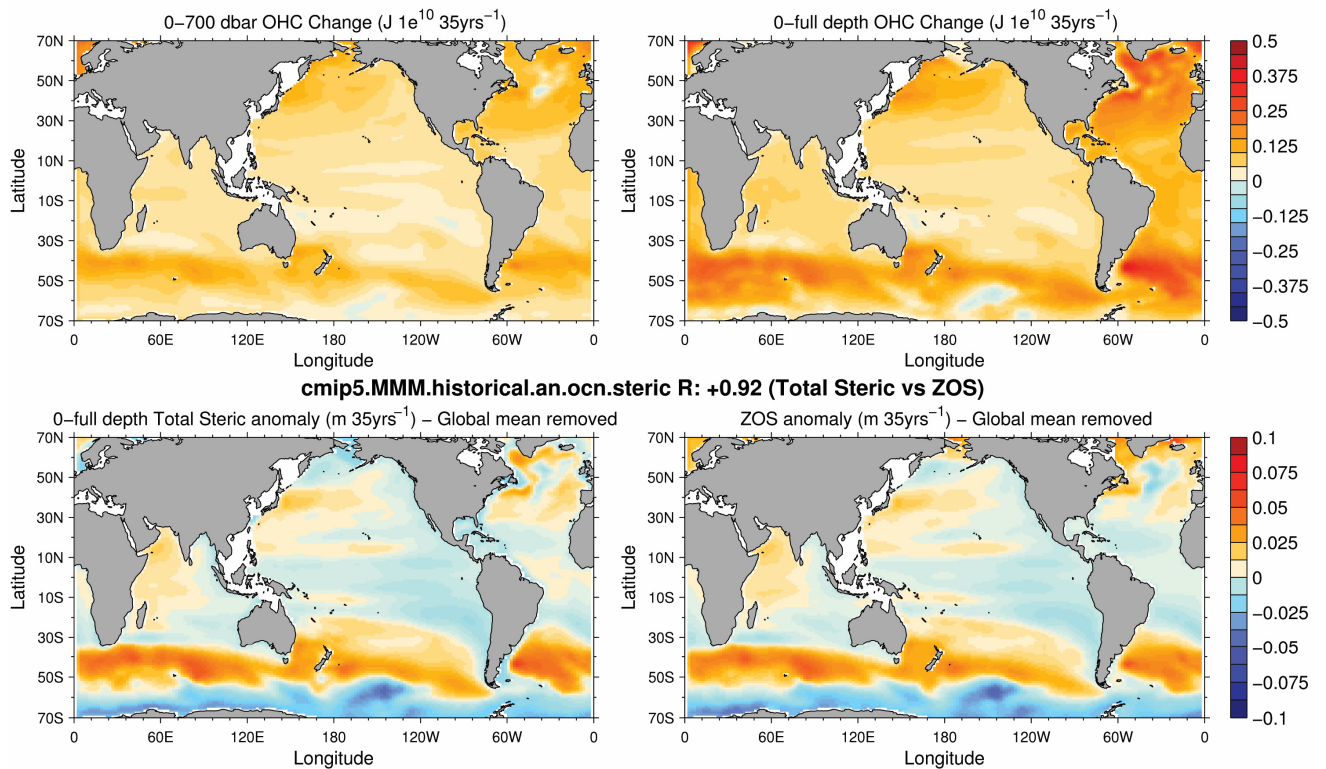


FIG. S5. 35-year multi-model mean (MMM) trends of ocean heat content (OHC) for the A) upper 0-700 dbar and B) 0-full depth ocean, C) 0-full depth total steric sea level (global mean removed) and D) sea surface height (SSH/CMIP5 variable ZOS; global mean removed).

We find a very high correlation between modelled SSH (global mean removed) and steric height (global mean removed) derived from modelled salinity and temperature fields. Discrepancies between our derived total steric fields and the model SSH fields (lower panels Figure S5) can be attributed to differing versions of the equation of state used across the models. Additionally, SSH reflects the full ocean depth, rather than the more focused 0-700 dbar analyses that are most common (see Supplemental Section 1), and consequently, deep ocean drifts can influence results (see discussion in Section 7). The MMM spatial correlation between total steric anomalies and SSH is 0.92; However, this number is calculated by excluding the GISS models, which show poor spatial

correlations of ~0.5 rather than the 0.89-0.99 shown from other models contributing to CMIP5.

The linear relationships and strong correlations shown in Figures 2, S4 and S5 support our rationale to use SSH, or – more specifically – rates of hemispheric sea-level changes, as a proxy from which to adjust SH OHC change estimates through the use of hemispheric ratios.

Sea surface height trends – observations and CMIP5 simulations

The SSH trend pattern observed over the last 20 years (1993 – 2012) features a prominent East-West dipole structure in the tropical Pacific Ocean (Figure S6). Much of this pattern has been attributed to internally forced interannual to decadal climate variability, mostly related to ENSO (interannual) and PDO-like (decadal) variability (e.g., Meyssignac *et al.*, 2012; Zhang & Church, 2012). Similarly, Hu & Deser (2013) have shown in CMIP5-type forced 21st century simulations, that internal variability will likely dominate the regional SSH trend pattern over many regions for the next 40-50 years, after which a forced global mean rise tends to emerge from the variability. However, they also show that the global mean OHC change varied only minimally between simulations in their 40-member ensemble. While this latter point was shown for the global mean, we posit here that it also holds over hemispheric averages, but it will break down at smaller regional scales due to the increasing dominating effect of internal variability that masks any externally forced SSH-trend mode.

Based on these previous findings and the uninitialised configuration of CMIP5 historical simulations, it is expected that the observed simulated SSH regional 20-year trend patterns (Fig S6A; global mean subtracted) would not likely be replicated by the CMIP5 MMM. This is indeed the case when we compare the single observed realization (Figure S6A) against the MMM 20-year trend (Figure S6B). While the time periods of the 20-year trends in Figure S6 only partially overlap (A: 1993-2012 vs B: 1985-2004), the conclusions are not sensitive to the particular choice of time period (Hu & Deser, 2013). The strong tropical Pacific di-pole signal in observed SSH trends is largely absent in the CMIP5 MMM. Although observed and simulated 20-year SSH trends agree in sign in some regions (Figure S6), we note that all grid points in the MMM SSH 20-year trend pattern have uncertainties that are larger than the trends themselves (stippled regions where less than 75% of models agree with the sign of the MMM trend). However, CMIP models are capable of reproducing the observed 20-year trend patterns in the Pacific (as demonstrated in Meyssignac *et al.*, 2012), but not in phase with observations due to the uninitialised configuration. Therefore, we stress again that a lack of strong agreement between the CMIP5 MMM and AVISO SSH trend patterns over a specific time period (such as 1993-2012 or 1985-2004) is entirely expected, and is not a measure of model performance. It is pertinent to distinguish between the ability of a model to simulate a long-term warming trend, and the ability of a model to match observed changes – in particular when the period of comparison is relatively short (i.e. internal variability tends to “mask” the long-term, externally forced trend).

While the phasing of non-synchronised internal variability limits a direct comparison of sea-level change patterns between CMIP simulations and observations, the comparison of hemispheric ratios of these rates circumvents this problem. Effectively, the ratios (as shown in Figure 3) normalize each hemispheric trend estimate by the global mean (sea level rise or OHC), and – in combination with spatial averaging – eliminate the phase-dimension from the comparison. The agreement of hemispheric SSH trend ratios among the CMIP models (Figure 3A, C) indicates that a common limit on hemispheric sea-level partitioning exists, irrespective of the presence of different internal variability in each model for the particular time period assessed. Therefore, we argue that because (1) the models' hemispheric SSH ratios agree with the observed ratios, and (2) the close correspondence between OHC and SSH at hemispheric and global scales exists (Figure 2, S4), the model-derived OHC hemispheric ratios can be used to adjust the poorly constrained observed Southern Hemisphere OHC change estimates.

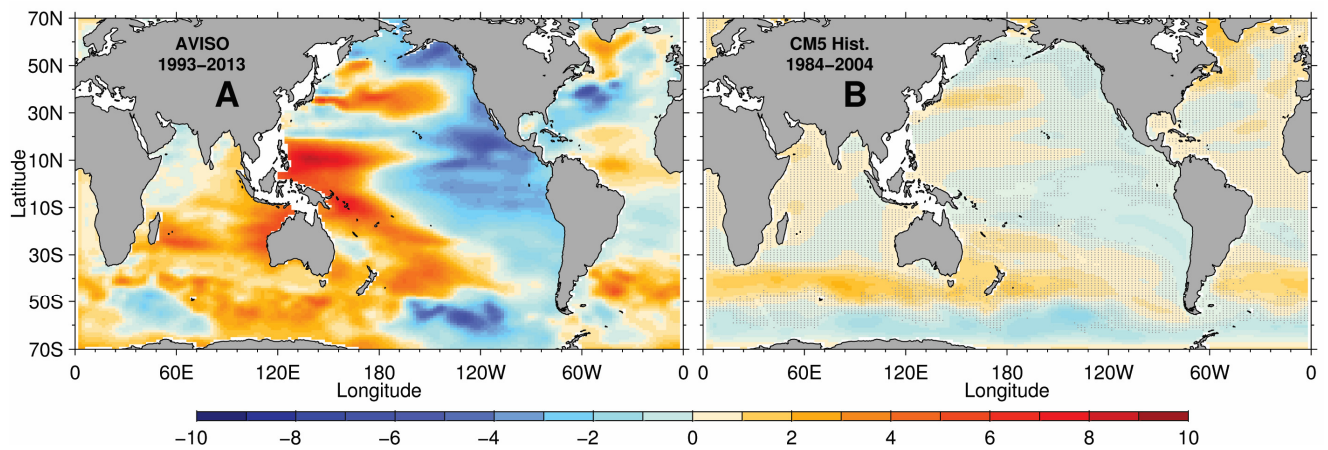


FIG. S6. 20-year trends (global mean subtracted) in A) AVISO SSH (1993-2012) and the B) CMIP5 Historical MMM (1985-2004); units are mm yr^{-1} . In B) the stippling indicates where less than 75% of models agree in the sign of the MMM – and that for almost all regional trends the uncertainty (largely due to internal variability) is larger than the MMM trend.

4. Signal and noise across CMIP5 experiments – unforced internal variability

Quantifying the role of unforced variability is an important aspect when considering long-term changes to OHC. We compare spatial pattern correlations (PCs) between the four observational OHC change estimates for which maps were available (Figure S2a), each model from the CMIP5 Historical experiment, and the most strongly forced future experiment (RCP85). The low correlations for any single model realisation (Figure S7: red diamonds) or a single model ensemble average compared to observations (Figure S7: grey bars) can result from observational deficiencies, structural errors/biases in the models, and sampling uncertainties associated with unforced internal climate variability. We estimate the importance of unforced internal climate variability by computing the PCs between pairs of Historical simulations from single models (Figure S7A, black diamonds). These PCs range from -0.34 to 0.73 (MMM of 0.31) suggesting that internal variability is large and plays an important role. The importance of internal variability is reduced in the RCP85-forced 21st century experiments because the externally forced response is significantly larger than those in Historical simulations (see Figure S2b). In these RCP85 simulations, the spatial patterns are more consistent across a single model's ensemble of simulations, with PCs that range from 0.41 to 0.97 (Figure 2B: black diamonds; MMM of 0.75). The fact that the more strongly forced RCP85 projections correlate better with the observations than the Historical simulations suggests a stationary, and increasingly evident model fingerprint of forced change (Figure S7A/B horizontal red [0.12 versus 0.21] and grey [0.15 versus 0.22] lines). Despite the obscuring role played by internal variability, the PCs are generally higher between pairs of model simulations than between models and observations (Figure S7A/B black line and red lines), even for Historical simulations. Overall, consistently low PCs in the Historical experiments imply that even with model improvements and large ensembles, increasing confidence in regional-scale projections of OHC continues to be a challenge.

Keeping in mind the deficiency of SH sampling, we further compare the observations and the Historical MMM by considering correlations for the NH and SH separately (Figure S7C & D). NH correlations are considerably higher, with an average inter-observation value of 0.68 compared to a SH value of 0.44 (Figure S7C & D; horizontal grey lines). This reflects the weak observational constraint due to poor data coverage in the SH, with infilling methods leading to pattern differences between the observational data sets. These results provide a motivation to further investigate the dissimilarity, and contrast the model-simulated and observed hemispheric partitioning of OHC changes.

The DW10 analysis was selected as the observed baseline, as it consistently obtained the highest mean PC compared to the MMM (Figure S7). We note the Smi07 analysis provided similar PC values. Qualitative investigations with the Ish09 and Lev12 analyses were undertaken and the larger PC of DW10 was attributed to its higher SH values (of similar magnitude to NH values). In these other analyses, equivalent PC values to DW10 were obtained when their Southern Hemisphere trends were artificially enhanced by 50%.

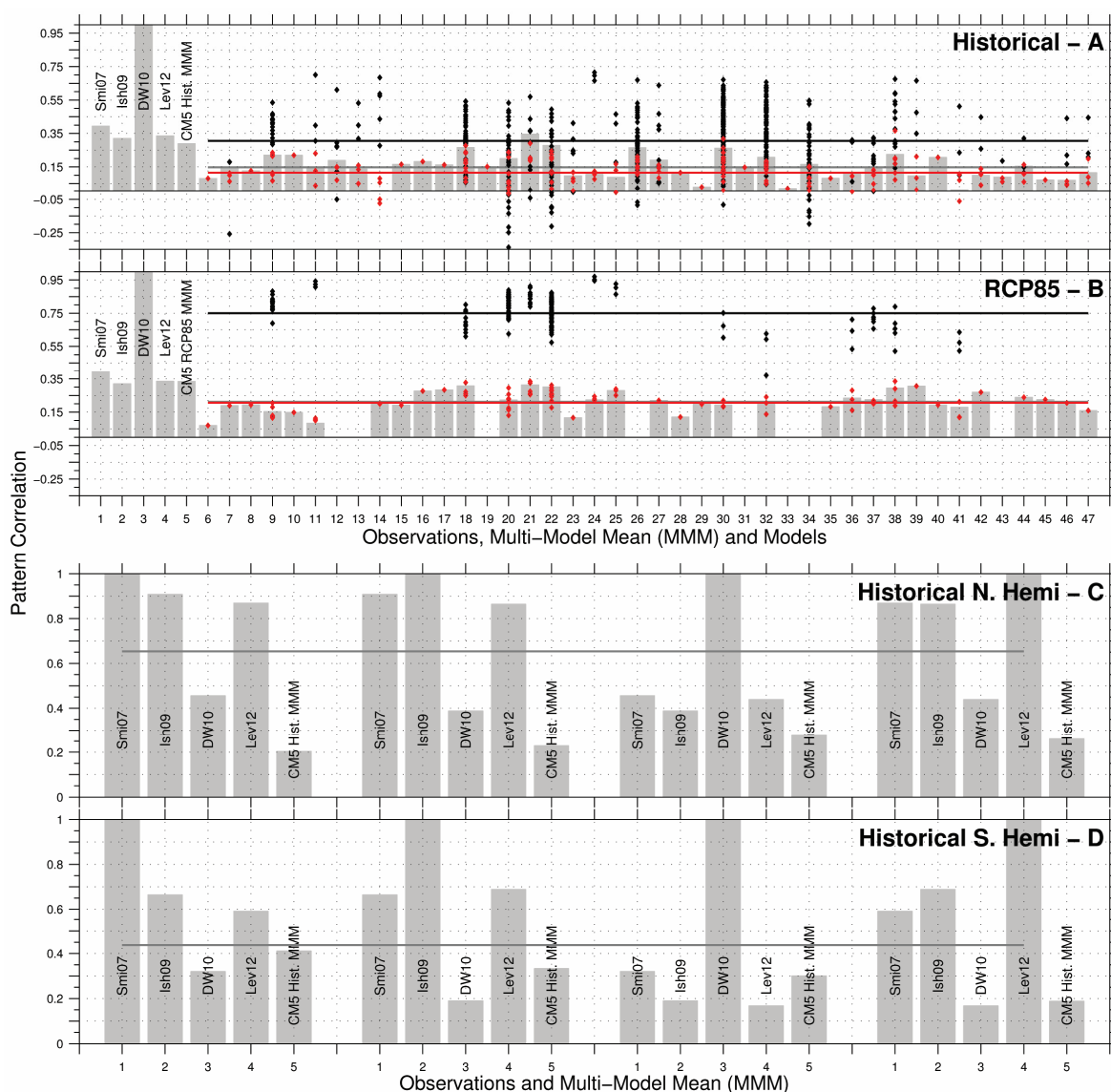


FIG. S7. Pattern correlation (PC) between 35-year trends maps of 0-700 dbar ocean heat content for 4 observational estimates (numbers 1-4), the CMIP5 multi-model mean (MMM; 5), and each of the available CMIP5 models (6-47) for the Historical 1970-2004 (A) and RCP85 2065-2099 (B) experiments (grey bars). The DW10 analysis (3) is the reference to which each of the observed analyses and model simulations is compared, as it consistently showed the highest mean PC compared to the MMM (5; panel A & B). Red diamonds indicate the PC between independent model simulations (i.e. r1i1p1) against observations, and black diamonds indicate the PC between pairs of model simulations within the model suite (i.e. r1i1p1 vs r2i1p1). The grey horizontal line expresses the average MMM PC value for each independent model ensemble (6-47; grey bars), the red horizontal line shows the weighted (each model equally contributes to the mean regardless of the number of simulations) MMM PC value for each independent model realisation (6-47; mean of red diamonds) and the black line expresses the MMM PC value from within the model suite (6-47; mean of black diamonds). A listing of models is contained in Table S1. Lower panels C & D express the Hemispheric inter-observation PC for each of the 4 observational estimates (Smith & Murphy [2007: 1] on the left hand side and Levitus *et al.* [2012: 4] on the right hand side; See Table S1) and the CMIP5 Historical MMM for the Northern (C) and Southern Hemisphere (D) respectively. The grey horizontal line expresses the average observational PC value across each independent observational combination (grey bars), excluding self-correlation (values of 1) and the CMIP5 Historical MMM.

5. Full CMIP ensemble supporting analysis

We investigate hemispheric ocean warming ratios in all CMIP3 and CMIP5 simulations to evaluate the role of forcing uncertainties and finite simulation samples. To achieve this we considered the full suite of results across 9 CMIP experiments (CMIP3: 20c3m, SRESA1B, SRESA2 & SRESB1; CMIP5: Historical, RCP26, RCP45, RCP60 & RCP85). The following results consider each independent simulation equally (in contrast to summary figures within the main text which have considered independent model ensemble means: averaging each available realisation for each independent model to provide a single model result), and consequently model pathologies, errors and biases present in single simulations affect the following results.

We find that the full ensemble results support those obtained in the main text, however these yield larger uncertainties due to outlier single simulations (we exclude any simulations that are outside the 99.9% confidence interval of the full 641 realisation suite in calculations, which reduces the realisation count to 633). We also find that the spread of results is larger in the more weakly forced simulations (SRESB1, RCP26, RCP45) and are even greater in simulations that include natural forcings such as volcanoes (20c3m, Historical). In contrast to results presented in Figure 4, the MMM averages of the SH versus the global mean OHC ratio for the full suite lie at 0.55 and 0.58 for the 20c3m/Historical composite (Figure S8).

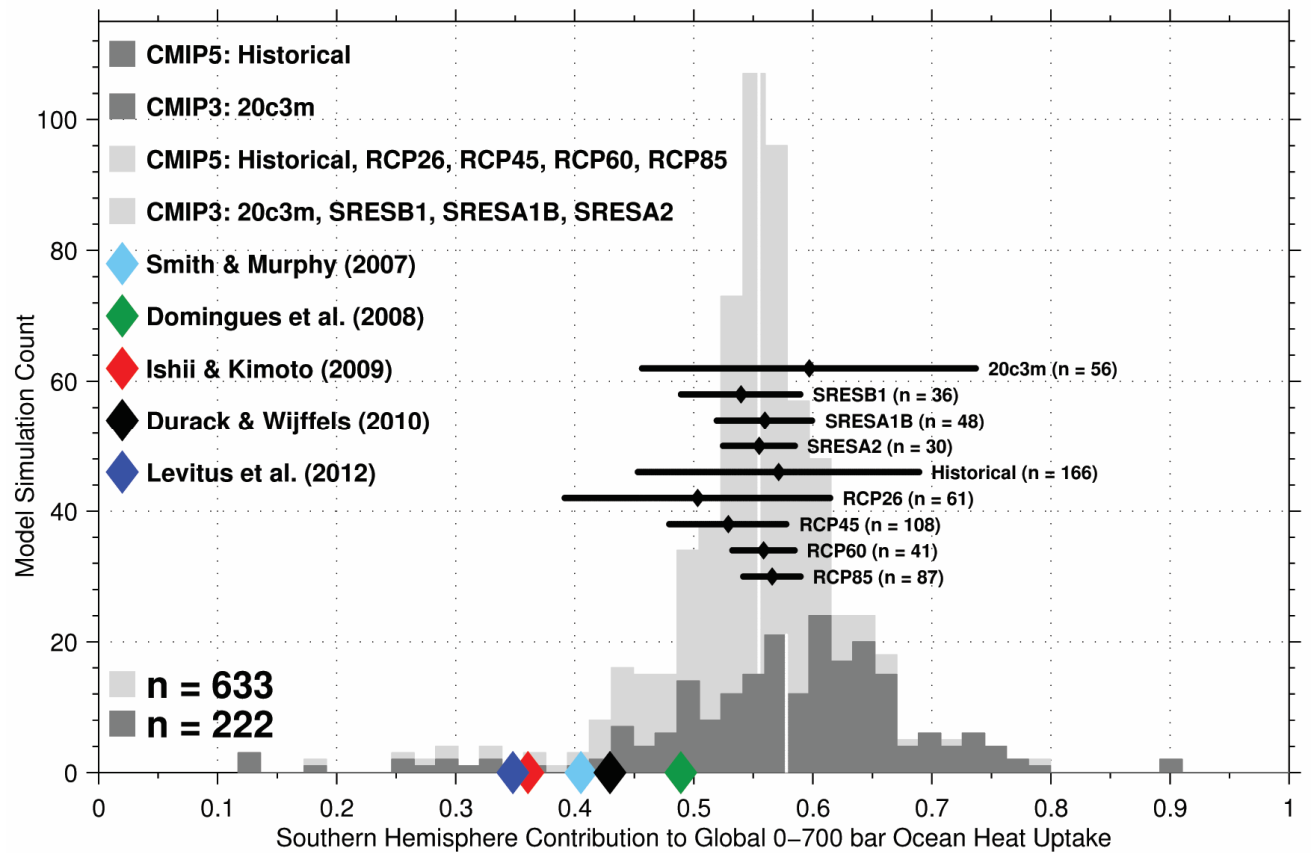


FIG. S8. Southern Hemisphere versus global 35-year OHC change as obtained from a large suite of available 20th and early 21st century dark grey and 20th and 21st century light grey simulations and 5 independent observational estimates of Smith & Murphy (2007; light blue diamond), Domingues *et al.* (2008; green diamond), Ishii & Kimoto (2009; red diamond), Durack & Wijffels (2010; black diamond) and Levitus *et al.* (2012; dark blue diamond). The full model distribution as expressed from 633 independent CMIP3 (20c3m, SRESB1, SRESA1B, SRESA2) and CMIP5 (Historical, RCP26, RCP45, RCP60, RCP85) simulations is expressed in light grey, and 222 20th and early 21st century 20c3m (CMIP3) and Historical (CMIP5) simulations (dark grey) that can be directly compared to observations. Horizontal black lines represent each of the CMIP experiments showing the multi-model mean (MMM; small black diamonds) and 1 standard deviation around the MMM (horizontal black line). Vertical white lines denote the ensemble mean for the full ensemble (light grey) and 20c3m/historical ensemble (dark grey) respectively. We note that all simulations (rather than single model-ensemble means) are used, which contrasts with the approach taken in Figure 4 and provides larger uncertainties.

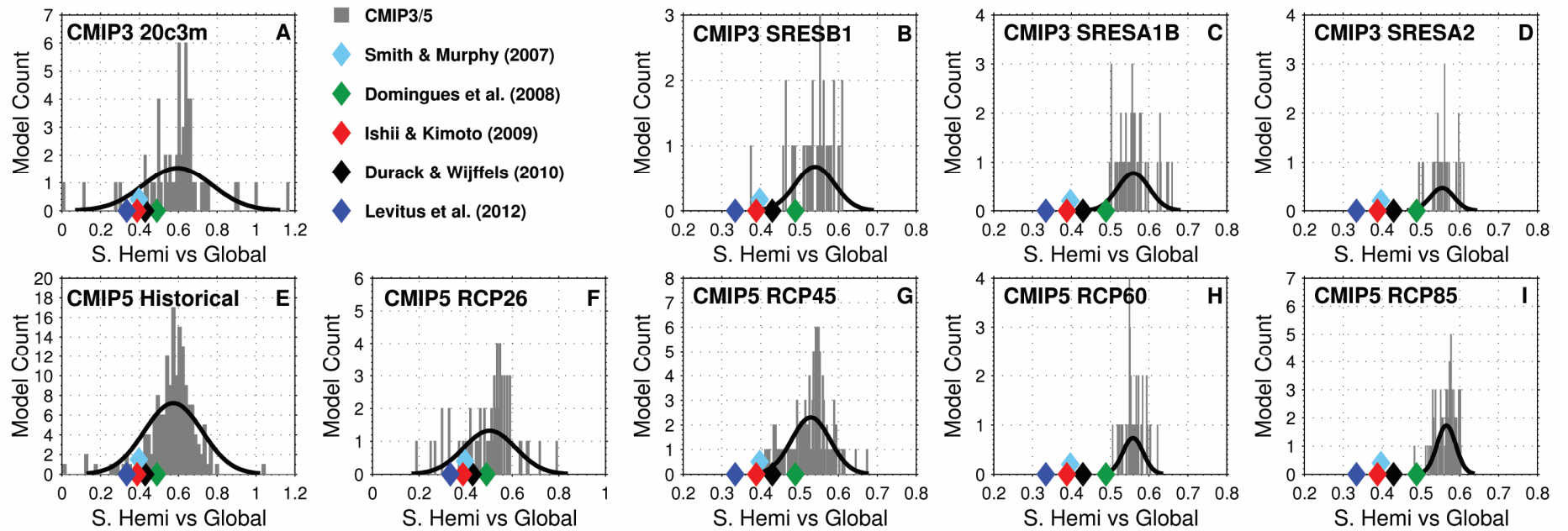


FIG. S9. Southern Hemisphere versus global 35-year heat content change as obtained from a large suite of available CMIP3 (20c3m, SRESB1, SRESA1B, SRESA2) and CMIP5 simulations (Historical, RCP26, RCP45, RCP60, RCP85) and 5 independent observational estimates of Smith & Murphy (2007; light blue diamond), Domingues *et al.* (2008; green diamond), Ishii & Kimoto (2009; red diamond), Durack & Wijffels (2010; black diamond) and Levitus *et al.* (2012; dark blue diamond).

6. Data sources

Models

The model data used in this study is a combination of data contributed to the CMIP5 (Taylor *et al.*, 2012) and CMIP3 (Meehl *et al.*, 2007) archives that represent the state of the art in coupled climate modelling for 2011 and 2005 respectively. The CMIP5 experiments analysed include: Historical (1970-2004; 42 models), RCP26, RCP45, RCP60 and RCP85 (2065-2099). The CMIP3 experiments analysed include: 20c3m (1970-1999; 23 models), SRESB1, SRESA1B and SRESA2 (2065-2099). Specific details on the model simulations are contained in Tables S1 (CMIP5) and S2 (CMIP3). The anthropogenic forcing information is detailed in Meinshausen *et al.* (2011: CMIP5) and Nakićenović and Swart (2000: CMIP3) with natural forcings for CMIP5 described in Vernier *et al.* (2011), Driscoll *et al.* (2012), Schmidt *et al.* (2012) and Eyring *et al.* (2013). Improvements in ocean simulation when comparing CMIP3 to CMIP5 include: increased model resolution, both in horizontal and vertical grids, along with better eddy parameterisations focusing on improving eddy fluxes and diffusivity (Flato *et al.*, 2013). An investigation of inter-model forcing differences and uncertainties for the CMIP5 model suite has been documented by Forster *et al.* (2013).

Observational OHC data

We use five observational analyses of long-term ocean temperature changes, from which the key observed ocean temperature change conclusions in the latest release of the IPCC Fifth Assessment Report (AR5) were made (Rhein *et al.*, 2013). These analyses are selected as they provide the most comprehensive assessment of different estimates of temperature changes, using differing methodologies, XBT bias correction schemes and their inherent assumptions (Abraham *et al.*, 2013).

The Palmer *et al.* (2007) analysis was excluded, as their technique does not provide infilled spatial maps, with their heat content timeseries dependent on the assumption that unsampled regions warm at the global average rate and for this reason could not be directly compared to other infilled products (see Supplemental Section 1).

The analysis of Smith & Murphy (2007; Smi07) uses covariances obtained from a coupled global climate model (HadCM3) to guide an optimal interpolation of monthly subsurface ocean temperature and salinity data from 1950 to 2013. The Domingues *et al.* (2008; Dom08) analysis uses a reduced-space optimal interpolation technique, which is optimized to recover broad-scale patterns from sparse spatial data from 1950 to 2008. The Dom08 gridded analysis was not available, however for this study we have used the basin results as presented in Gleckler *et al.* (2012). The analyses of Ishii & Kimoto (2009, updated v6.13; Ish09) and Levitus *et al.* (2012; Lev12) use an objective-mapping technique to generate annual mean gridded temperature maps for the period 1945-2012 and 1955-2012 respectively. For the Lev12 analysis, we used the corresponding temperature and salinity climatology from the World Ocean Atlas 2009 (WOA09). The analysis of Durack & Wijffels (2010; DW10) uses a spatial and temporal parametric

model, optimized to recover the broad-scale ocean mean structure, the annual and semi-annual cycle (and their spatial gradients) and the multidecadal linear trends from the sparse hydrographic database (salinity and temperature profiles from bottles, CTDs and Argo) over 1950-2008, a dataset free of known biases. A key advantage of the DW10 analysis is the exclusion of XBT profile data (the original analysis was generated to ascertain long-term salinity changes), which have been isolated as the source of the biases in the thermal archive (Gouretski & Koltermann, 2007), although the exclusion of XBT data is also a disadvantage with less than half the profiles contained in the thermal archive (no salinity) alone (also see Section 1).

Observational SSH data

We use the monthly gridded AVISO reference sea surface height (SSH/MSLA) data (Ducet *et al.*, 2000), which is a merged satellite product that homogenises data from the Topex/Poseidon, Envisat, Jason-1 and Jason-2 satellites. For the analysis of 20-year trends, we use the data span from January 1993 through December 2012, and regridded the AVISO fields onto a regular 1 x 1 degree (longitude, latitude) grid. To compute spatial averages, we use the same land-sea mask as the CMIP ensemble.

7. Data preparation and analysis

For all model data and observational estimates for which monthly gridded data was available (Smi07, Ish09, AVISO), annual means were first generated.

Ocean heat content maps

For all data series a linear least-squares trend and climatology was calculated at each grid cell for 1970-2004 on the native observed analysis or model grid. For DW10, spatial maps presenting the linear trend and climatology over 1970-2008 were generated. All trend maps and climatologies were then interpolated to a 2 x 1 (longitude, latitude) degree grid extending from 70°S to 70°N which excludes marginal seas (Mediterranean Sea, Baltic Sea, Red Sea, Persian Gulf, China Seas, Sea of Japan, Java Sea, Banda Sea and Arafura Sea) and vertically interpolated to 18 standard pressure levels (5, 10, 20, 30, 40, 50, 75, 100, 125, 150, 200, 300, 500, 700, 1000, 1500, 1800, 2000 dbar) for direct comparison. To compare observations and models fairly (which have differing land-sea masks), after interpolation of results to the target grid an iterative nearest neighbour infilling algorithm was employed to infill regions so that the land-sea masks of all analyses were identical. Heat content anomalies are calculated using observed or modelled temperature, salinity and pressure along with derived specific heat capacity and in-situ density which are calculated from the corresponding climatological mean. Interpolated maps provide results presented in Figures 1-2, 4-5, S2a, S2b, S4, S5, S6, S7, S8, S9, S10 and S11.

Ocean heat content timeseries

We used annual means of the sea water potential temperature ('THETAO') fields from CMIP5 Historical simulations to analyse and compare hemispheric ocean heat content

(OHC) trends against the 4 observed estimates of OHC change. The processing steps taken to analyse these fields involved:

1. Integrating each annual modelled 'THETA0' or observed temperature field over depths from 0-700 m using the native model or observational vertical coordinate (upper most grid cell to 700 m). When integrating over depth, each value is weighted by the cell depth so a depth-weighted value is obtained for each grid cell.
2. Integrating horizontally for the Southern Hemisphere (SH) and Northern Hemisphere (NH) separately using the native model or observational horizontal grid. When summing over grid cells, each value is weighted by the cell area so an area-weighted value is obtained for each grid cell. This calculation yields a quantity proportional to heat content for each hemisphere.
3. From the hemispheric OHC timeseries; calculating linear least-square trends of length 1 - 35 years for models and 1 to 43 years for observations, starting in 1971 and contained in the 1971 – 2004 interval for models (2012 for observations)
4. Calculating the trend ratio for OHC trends, SH to global (i.e. SH/(NH+SH))
5. Plotting observed and MMM ratios as a function of trend length, and also plotting the 1 standard deviation spread across models

Hemispheric OHC timeseries provide results presented in Figure 3B, D.

Sea surface height

We used annual means of the 'ZOS' (dynamic sea surface height above geoid) and 'ZOSTOGA' (global average thermosteric sea level change) fields from CMIP5 Historical simulations to analyse and compare hemispheric SSH trends against the AVISO satellite altimetry observations. The processing steps taken to analyse these fields involved:

1. The global mean at each timestep of the 'ZOS' fields is zero, in accordance with the model definition of dynamic SSH; all 'ZOS' fields were regridded to a 1 x 1 (longitude, latitude) 70°S to 70°N grid.
2. For each available CMIP5 realisation, we computed and subtracted the 'ZOSTOGA' preindustrial control (piControl) run drift from the Historical realisation's 'ZOSTOGA' field; drift was computed as a linear least-squares fit over the concurrent time period of the piControl and Historical simulations, i.e. from the time period the Historical run branches off from the source piControl (typically between years 1850-1870) through 2004.
3. Finally, 'ZOS' and the adjusted 'ZOSTOGA' fields were added to obtain the full dynamic and steric sea level change fields. Linear least-square trends were then calculated over varying time-windows as described below and averaged with equal-area weighting over the same surface areas as the OHC analysis. If global mean SSH values were less than 0.5mm, this was reset to equal 0.5mm to prevent model ratios being overly inflated due to divide by a near-zero error.

Further details regarding the processing of the 'ZOS' and 'ZOSTOGA' fields in CMIP5 simulations can be found in Yin *et al.* (2013) and Landerer *et al.* (2013). To compare equal-weighted SSH trends against the volume-integrated OHC trend ratios, we scaled the Southern and Northern Hemisphere rates by 0.62 and 0.38 respectively, to reflect the different volumes for the Southern and Northern Hemisphere ocean regions. In this

way, a value of 0.62 for the Southern Hemisphere/Global SSH trends implies that SSH trends in the two hemispheres are equal. Hemispheric SSH timeseries provide results presented in Figure 3A, C.

Period of analysis

Previous work has highlighted discrepancies with observational analyses prior to 1970 (e.g. Gleckler *et al.*, 2012), so to fairly compare models and observations the 35-year period 1970-2004 was selected which allows direct trend comparison between all but one observational estimate (DW10) and the Historical simulations of the CMIP5 models. Splicing of the Historical and RCP simulations to generate 1970-2012 comparable timeseries was not undertaken, as this would have considerably reduced the number of individual simulations available for analysis. The DW10 analysis is dependent on the Argo period (~2003-2008) to resolve the oceans mean spatial structure, and so trends obtained over the 39-year (1970-2008) analysis were scaled to represent 35-years. To test for robustness of the forced spatial patterns, future RCP (CMIP5) and SRES (CMIP3) simulations over 2065-2099 are also assessed and presented alongside results obtained from the Historical (CMIP5; and 20c3m, CMIP3) experiment. For the CMIP3 models, the 30-year period 1970-1999 was selected (and trends scaled to represent an equivalent 35-year change) as the 20th century simulations terminate in 2000.

Multi-model mean (MMM) calculation

When multiple model simulations were available, a model-ensemble mean was generated. In plots presenting the multi-model mean (MMM) result, the trend maps from each model (or model-ensemble if available) were averaged to provide a single MMM map. Averaging multiple model simulations first before calculating the MMM ensures that each model contributes equally to the MMM, or to the resolved distribution (Figures 3, 4, 5, S2a, S2b, S3, S5, S6 and S7). In distribution plots for which each independent realisation is sampled (Figure S8, S9, S10 and S11) each single realisation is equally weighted making the assumption that each independent model realisation is an equally-likely alternative sample.

To prevent outlier simulations from affecting ensemble statistics when using the full model suite (all experiments presented within each figure), we exclude single model-ensemble means (Figure 4) or single simulations (Figure S8 and S11) that exceed the 99.9% confidence intervals obtained from the suite. This led to a reduction in model-ensemble means contributing to Figure 4 (CMIP3 20c3m: 2; CMIP5 Historical: 1; and experiment bounds [black horizontal lines] CMIP5 RCP26: 1) and to single simulations contributing to Figure S8 (CMIP3 20c3m: 4; CMIP5 Historical: 4) and Figure S11 (CMIP5 Historical: 1). This in effect reduced the uncertainty presented for experiment spreads (Figures 4 and S8) and the ensemble distribution (Figure S11). Prior to the 99.9% confidence intervals being calculated for Figure S8, the HadGEM2-CC.historical.r2i1p1 realisation was removed. This was due to the single simulation having an order of magnitude larger hemispheric ratio when compared to the full model suite. The outlier affected the calculation of statistics and so was removed.

Model drift

The effect of model drift (e.g. Rahmstorf, 1995; Covey *et al.*, 2006; Gleckler *et al.*, 2012; Sen Gupta *et al.*, 2012, 2013), which is a particularly relevant issue for the deeper ocean, was assessed. To investigate the role of drift a contemporary 150-year portion of the piControl simulation (1900-2049) was analysed for the historical simulations in which data was available, and a linear fit to this timeseries was undertaken at each grid point in three dimensions. This linear drift estimate was then subtracted grid point by grid point from the corresponding linear historical trend estimate on the model native grid before calculation of heat content change was undertaken. A sensitivity study was undertaken to ascertain how accounting for drift can change the Southern Hemisphere to Global OHC change ratio, and found that hemispheric ratios are largely insensitive to drift correction when considering the pooled ensemble average (Figure S10). It was found that accounting for drift reduced the CMIP5 Historical MMM Southern Hemisphere to global OHC change ratio by less than 2% (Figure S11), and led to a slightly larger spread when comparing Southern Hemisphere and Global OHC for this single experiment (Figure S11). This suggests that the key results of the study are insensitive to model drift. Specific details on the model simulations used in this analysis are contained in Tables S1 & S2.

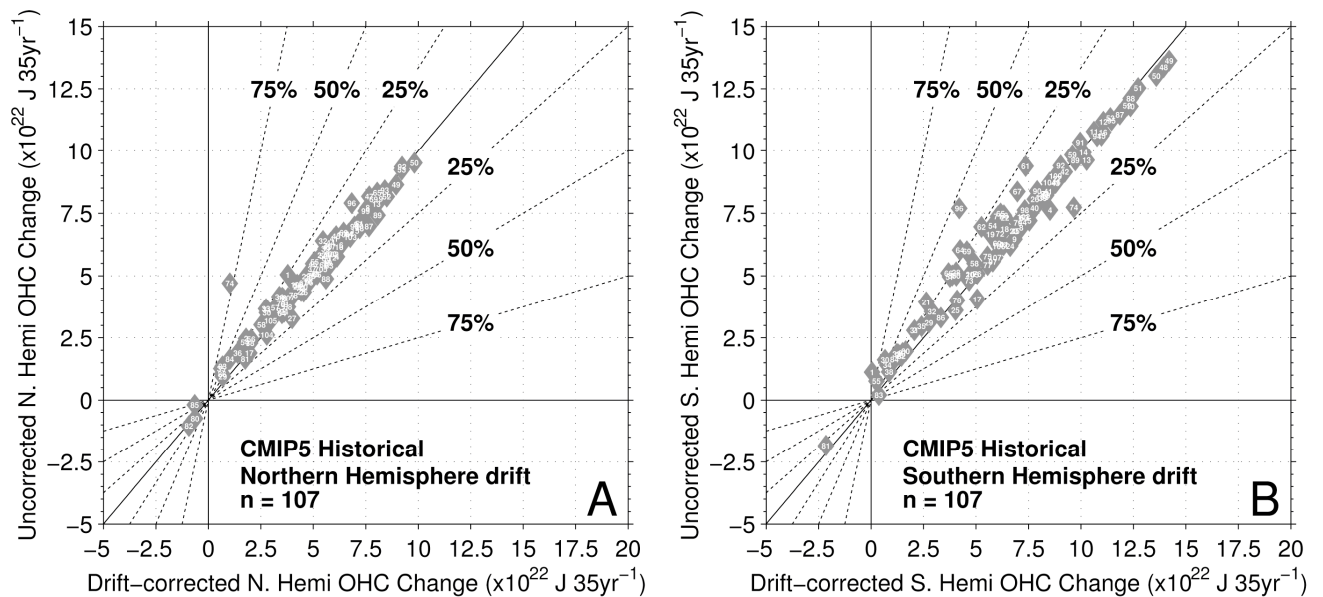


FIG. S10. The effect of drift-correction on 0-700 dbar ocean heat content (OHC) change for available CMIP5 historical simulations for the Northern Hemisphere (A) and Southern Hemisphere (B) integrated-totals respectively. Dotted lines show percentage differences from the drift-uncorrected results. Drift correction has a proportionally larger influence on Southern Hemisphere values, however most of these corrections have a small to negligible effect.

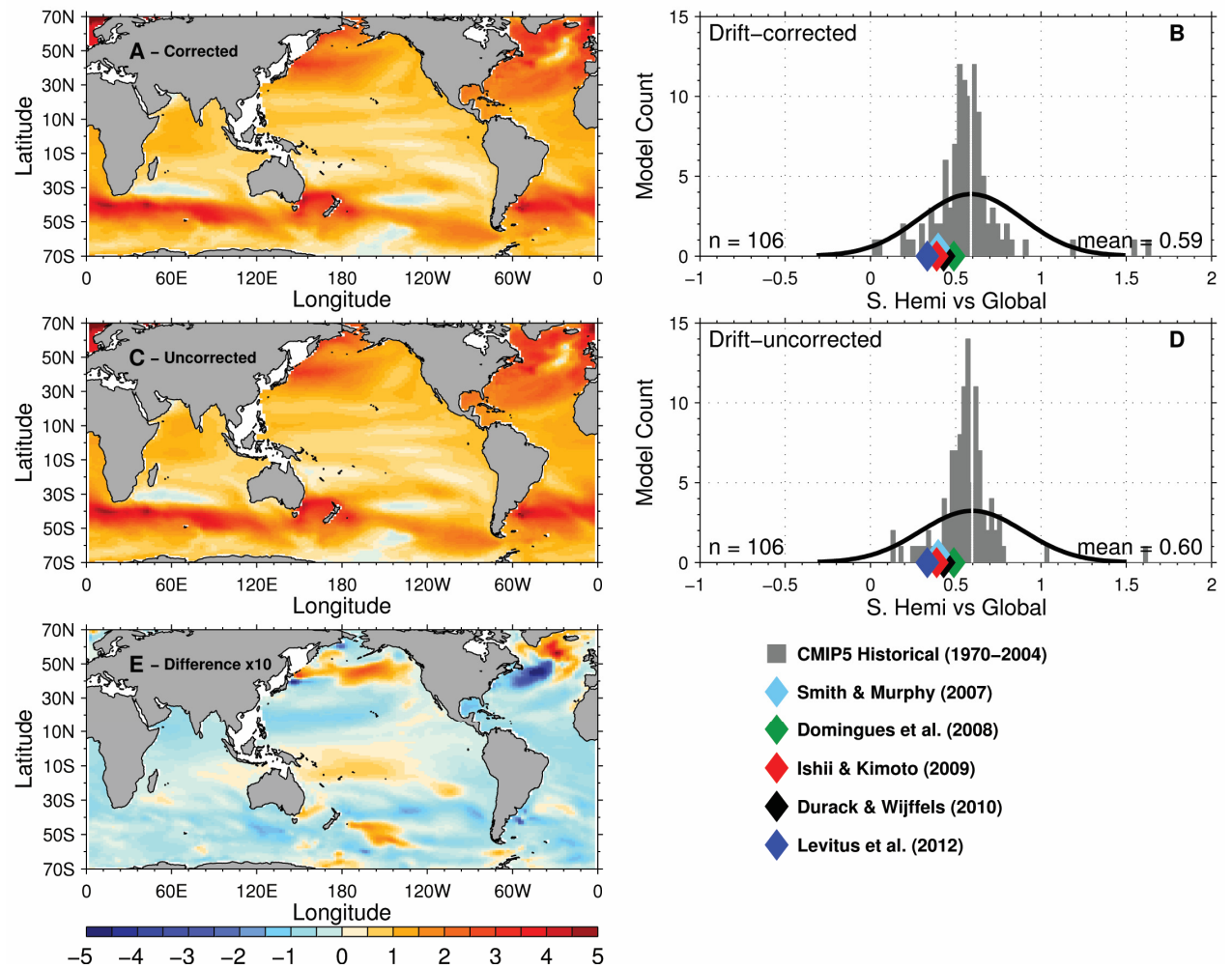


FIG. S11. The effect of drift-correction on 0-700 dbar ocean heat content (OHC) change for available CMIP5 Historical simulations. (A) drift-corrected multi-model mean (MMM) and (C) drift-uncorrected MMM maps and (B) drift-corrected and drift-uncorrected (D) Southern Hemisphere versus globally-integrated OHC change realisation distributions. The difference between the drift-corrected (A) minus drift-uncorrected MMM maps (C) is shown enhanced by 10x to show spatial patterns (E).

8. Supplementary tables

TABLE S1. Observational and CMIP5 model datasets

#	Model/Observation	OHC Models					SSH Models	
		1970-2004	2065-2099				Hist. drift correction	1970-2004
		Historical (Obs. version)	RCP26	RCP45	RCP60	RCP85	piControl	Historical
1	Smith & Murphy (2007) – 1970-2004	N/A						
	Domingues <i>et al.</i> (2008) – 1970-2004	2.0						
2	Ishii & Kimoto (2009) – 1970-2004	6.13						
3	Durack & Wijffels (2010) – 1970-2008	1.0						
4	Levitus <i>et al.</i> (2012) – 1970-2004	29 th April 2013						
-	AVISO – 1993-2012	-						
5	CMIP5 MMM							
6	ACCESS1-0	1	-	1	-	1	1	1-2
7	ACCESS1-3	1-3	-	1	-	1	1	1-3
8	BNU-ESM	1	1	1	-	1	-	-
9	CCSM4	1-6	1-6	1-6	1-6	1-6	1	1-6
10	CESM1-BGC	1	-	1	-	1	1	1
11	CESM1-CAM5	1-3	1-3	1-3	1-3	1-3	-	1-3
12	CESM1-CAM5-1-FV2	1-4	-	-	-	-	-	-
13	CESM1-FASTCHEM	1-3	-	-	-	-	-	1-3
14	CESM1-WACCM	1-4	2	2	-	2	-	1-4
15	CMCC-CESM	1	-	-	-	1	1	1
16	CMCC-CM	1	-	1	-	1	-	1
17	CMCC-CMS	1	-	1	-	1	1	1
18	CNRM-CM5	1-10	1	1	-	1,2,4,6,10	1	1-10
19	CNRM-CM5-2	1	-	-	-	-	-	1
20	CSIRO-Mk3.6.0	1-10	1-10	1-10	1-10	1-10	1	1-10
21	CanESM2	1-5	1-5	1-5	-	1-5	1	1-5
22	EC-EARTH	2,3,5-7,9,10,12,14	8,12	1-3,6-14	-	1-3,6-14	1	2-3,5-7,9-12,14
23	FGOALS-g2	1-3,5	1	-	-	1	-	3,5
24	FGOALS-s2	1-3	1	-	1	1-3	1	1-3
25	FIO-ESM	1-3	1-3	-	1-3	1-3	1	1-3
26	GFDL-CM2p1	1-10	-	-	-	-	-	1-10
27	GFDL-CM3	1-5	1	1,3,5	-	1	1	-
28	GFDL-ESM2G	1	1	1	1	1	-	1
29	GFDL-ESM2M	1	1	1	1	1	-	1
30	GISS-E2-H	1-6 (p=1-3)	1 (p=1-3)	1-5 (p=1-3)	1 (p=1-3)	1 (p=1-3)	1 (p=1-3)	-
31	GISS-E2-H-CC	1	-	1	-	1	1	-
32	GISS-E2-R	1-6 (p=1,2,3,121,122,124-128)	1 (p=1-3)	1-6 (p=1-3)	1 (p=1-3)	1 (p=1-3)	1	1-6 (p=1-3,121-122,124-128)
33	GISS-E2-R-CC	1	-	1	-	-	1	1
34	HadCM3	1-10	-	-	-	-	-	-
35	HadGEM2-AO	1	1	1	1	1	-	-
36	HadGEM2-CC	1-3	-	1	-	1-3	1	1-3
37	HadGEM2-ES	1-5	1-4	1-4	1-4	1-4	1	1-5
38	IPSL-CM5A-LR	1-6	1-4	1-4	1	1-4	1	1-6
39	IPSL-CM5A-MR	1-3	1	1	-	1	1	1-3
40	IPSL-CM5B-LR	1	-	1	-	1	1	1
41	MPI-ESM-LR	1-3	1-3	1-3	-	1-3	1	1-3
42	MPI-ESM-MR	1-3	1	1-3	-	1	1	1-3
43	MPI-ESM-P	1-2	-	-	-	-	1	1-2
44	NorESM1-M	1-3	1	1	1	1	1	1-3
45	NorESM1-ME	1	1	1	1	1	1	1
46	bcc-csm1-1	1-3	1	1	1	1	1	1-3
47	bcc-csm1-1-m	1-3	1	1	1	1	-	1-3
-	CanCM4	-	-	-	-	-	-	1-10
-	MIROC4h	-	-	-	-	-	1	1-3
-	MIROC5	-	-	-	-	-	1	1-5
-	MIROC-ESM	-	-	-	-	-	-	2
-	MIROC-ESM-CHEM	-	-	-	-	-	1	1

583 **TABLE S2. CMIP3 model datasets**

#	Model	OHC Models			
		1970-1999	2065-2099		
		20c3m	SRESB1	SRESA1B	SRESA2
	CMIP3 MMM				
1	bccr_bcm2_0	1	1	1	1
2	cccma_cgcm3_1	1-5	1-5	1-5	1-5
3	cccma_cgcm3_1_t63	1	1	1	-
4	cnrm_cm3	1	1	1	1
5	csiro_mk3_0	1-3	1	1	1
6	csiro_mk3_5	1-3	1	-	1
7	gfdl_cm2_0	1	1	1	1
8	gfdl_cm2_1	2	-	1	1
9	giss_aom	1-2	1-2	1-2	-
10	giss_model_e_h	1-5	-	1-3	-
11	giss_model_e_r	1-9	1	1-5	1
12	iap_fgoals1_0_g	1-3	1-3	1-3	-
13	ingv_echam4	1	-	1	1
14	ipsl_cm4	1-2	-	1	-
15	miroc3_2_hires	1	1	1	-
16	miroc3_2_medres	1-3	1-3	1-3	1-3
17	miub_echo_g	1-3	1-3	1-3	1-3
18	mpi_echam5	1-3	1,3	1-2	1
19	mri_cgcm2_3_2a	1-5	1-5	1-5	2-5
20	ncar_ccsm3_0	1,3	1-3	2,5,8	1-2,4
21	ncap_pcm1	3-4	-	2-4	4
22	ukmo_hadcm3	1-2	1	1	1
23	ukmo_hadgem1	1	-	-	-
-	inmcm3_0	-	1	1	1

584

9. Supplementary references

- Abraham, J. P., M. Baringer, N. L. Bindoff, T. Boyer, L. J. Cheng, J. A. Church, J. L. Conroy, C. M. Domingues, J. T. Fasullo, J. Gilson, G. Goni, S. A. Good, J. M. Gorman, V. Gouretski, M. Ishii, G. C. Johnson, S. Kizu, J. M. Lyman, A. M. MacDonald, W. J. Minkowycz, S. E. Moffitt, M. D. Palmer, A. R. Piola, F. Reseghetti, K. Schuckmann, K. E. Trenberth, I. Velicogna and J. K. Willis (2013) A Review of Global Ocean Temperature Observations: Implications for Ocean Heat Content Estimates and Climate Change. *Reviews of Geophysics*, **51** (3), pp 450-483. doi: 10.1002/rog.20022
- Abram, N. J., R. Mulvaney, F. Vimeux, S. J. Phipps, J. Turner and M. H. England (2014) Evolution of the Southern Annular Mode during the past millennium. *Nature Climate Change*, **4**, pp 564-569. doi: 10.1038/nclimate2235
- AchutaRao, K. M., B. D. Santer, P. J. Gleckler, K. E. Taylor, D. W. Pierce, T. P. Barnett and T. M. L. Wigley (2006) Variability of ocean heat uptake: Reconciling observations and models. *Journal of Geophysical Research: Oceans*, **111** (C5). doi: 10.1029/2005JC003136
- AchutaRao, K. M., M. Ishii, B. D. Santer, P. J. Gleckler, K. E. Taylor, T. P. Barnett, D. W. Pierce, R. J. Stouffer and T. M. L. Wigley (2007) Simulated and observed variability in ocean temperature and heat content. *Proceedings of the National Academy of Sciences of the United States of America*, **104** (26), pp 10768-10773. doi: 10.1073/pnas.0611375104
- Barnett, T. P., D. W. Pierce and R. Schnur (2001) Detection of Anthropogenic Climate Change in the World's Oceans. *Science*, **292** (5515), pp 270-274. doi: 10.1126/science.1058304
- Barnett, T. P., D. W. Pierce, K. M. AchutaRao, P. J. Gleckler, B. D. Santer, J. M. Gregory and W. M. Washington (2005) Penetration of Human-Induced Warming into the World's Oceans. *Science*, **309** (5732), pp 284-287. doi: 10.1126/science.1112418
- Boening, C., J.K. Willis, F.W. Landerer, R.S. Nerem and J. Fasullo (2012) The 2011 La Nina: So strong, the oceans fell. *Geophysical Research Letters*, **39** (19), L19602. doi: 10.1029/2012GL053055
- Bryan, F. O., G. Danabasoglu, P. R. Gent and K. Lindsay (2006) Changes in ocean ventilation during the 21st Century in the CCSM3. *Ocean Modelling*, **15** (3-4), pp 141-156. doi: 10.1016/j.ocemod.2006.01.002
- Bullister, J. L. (1989) Chlorofluorocarbons as Time-Dependent Tracers in the Ocean. *Oceanography*, **2** (2), pp 12-17. doi: 10.5670/oceanog.1989.03
- Bullister, J. L. and T. Tanhua (2010) Sampling and Measurement of Chlorofluorocarbons and Sulfur Hexafluoride in Seawater. In: *The GO-SHIP Repeat Hydrography Manual: A Collection of Expert Reports and Guidelines*. Hood, E.M., C.L. Sabine and B.M. Sloyan (Eds.). IOCCP Report Number 14, ICPO Publication Series Number 134. Available online at: <http://www.go-ship.org/HydroMan.html>

- Bullister, J. L. (2014) Updated (2014) Atmospheric CFC-11, CFC-12, CFC-113, CC14 and SF6 Histories. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Department of Energy, Oak Ridge, Tennessee. doi: 10.3334/CDIAC/otg.CFC_ATM_Hist_2014
- Capotondi, A., M. A. Alexander, N. A. Bond, E. N. Curchitser and J. D. Scott (2012) Enhanced upper ocean stratification with climate change in the CMIP3 models. *Journal of Geophysical Research*, **117**, C04031. doi: 10.1029/2011JC007409
- Cheng, L. and J., Zhu (2014) Uncertainties of the Ocean Heat Content Estimation Induced by Insufficient Vertical Resolution of Historical Ocean Subsurface Observations. *Journal of Atmospheric and Oceanic Technology*, **31** (6), pp 1383-1396. doi: 10.1175/JTECH-D-13-00220.1
- Church, J. A., N. J. White, R. Coleman, K. Lambeck and J. X. Mitrovica (2004) Estimates of the Regional Distribution of Sea Level Rise over the 1950-2000 Period. *Journal of Climate*, **17** (13), pp 2609-2625. doi: 10.1175/1520-0442(2004)017<2609:EOTRDO>2.0.CO;2
- Church, J. A., N. J. White, L. F. Konikow, C. M. Domingues, J. G. Cogley, E. Rignot, J. M. Gregory, M. R. van den Broeke, A. J. Monaghan and I. Velicogna (2011) Revisiting the Earth's sea-level and energy budgets from 1961-2008. *Geophysical Research Letters*, **38** (18), L18601. doi: 10.1029/2011GL048794
- Church, J. A., D. Monselesan, J. M. Gregory and B. Marzeion (2013) Evaluating the ability of process based models to project sea-level change. *Environmental Research Letters*, **8** (1), 014051. doi: 10.1088/1748-9326/8/1/014051
- Covey, C., P. J. Gleckler, T. J. Phillips and D. C. Bader (2006) Secular trends and climate drift in coupled ocean-atmosphere general circulation models. *Journal of Geophysical Research*, **111**, D03107. doi: 10.1029/2005JD006009
- Cowley, R., S. Wijffels, L. Cheng, T. Boyer and S. Kizu (2013) Biases in Expendable Bathythermograph Data: A New View Based in Historical Side-by-Side Comparisons. *Journal of Atmospheric and Oceanic Technology*, **30** (6), pp 1195-1225. doi: 10.1175/JTECH-D-12-00127.1
- Danabasoglu, G., S. Peacock, K. Lindsay and D. Tsumume (2009) Sensitivity if CFC-11 uptake to physical initial conditions and varying surface forcing in a global ocean model. *Ocean Modelling*, **29** (1), pp 58-65. doi: 10.1016/j.ocemod.2009.02.011
- Dixon, K. W., J. L. Bullister, R. H. Gammon and R. J. Stouffer (1996) Examining a coupled climate model using CFC-11 as an ocean tracer. *Geophysical Research Letters*, **23** (15), pp 1957-1960. doi: 10.1029/96GL01470
- Domingues, C. M., J. A. Church, N. J. White, P. J. Gleckler, S. E. Wijffels, P. M. Barker and J. R. Dunn (2008) Improved estimates of upper-ocean warming and multi-decadal sea-level rise. *Nature*, **453** (7198), pp 1090-1093. doi: 10.1038/nature07080
- Downes, S. M., N. L. Bindoff and S. R. Rintoul (2010) Changes in the Subduction of Southern Ocean Water Masses at the End of the Twenty-First Century in Eight IPCC Models. *Journal of Climate*, **23** (24), pp 6526-6541. doi: 10.1175/2010JCLI3620.1

- Downes, S. M. and A. McC. Hogg (2013) Southern Ocean Circulation and Eddy Compensation in CMIP5 Models. *Journal of Climate*, **26** (18), pp 7198-7220. doi: 10.1175/JCLI-D-12-00504.1
- Driscoll, S., A. Bozzo, L. J. Gray, A. Robock and G. Stenchikov (2012) Coupled Model Intercomparison Project 5 (CMIP5) simulations of climate following volcanic eruptions. *Journal of Geophysical Research*, **117** (D17105). doi: 10.1029/2012JD017607
- Durack, P. J. and S.E. Wijffels (2010) Fifty-Year Trends in Global Ocean Salinities and Their Relationship to Broadscale Warming. *Journal of Climate*, **23** (16), pp 4342-4362. doi: 10.1175/2010JCLI3377.1
- Ducet, N., P. Y. Le Traon and G. Reverdin (2000) Global high-resolution mapping of ocean circulation from TOPEX/Poseidon and ERS-1 and -2. *Journal of Geophysical Research*, **105** (C8), pp 19477-19498. doi: 10.1029/2000JC900063
- Dutay, J.-C., J. L. Bullister, S. C. Doney, J. C. Orr, R. Najjar, K. Caldeira, J.-M. Campin, H. Drange, M. Follows, Y. Gao, N. Gruber, M. W. Hecht, A. Ishida, F. Joos, K. Lindsay, G. Madec, E. Maier-Reimer, J. C. Marshall, R. J. Matear, P. Monfray, A. Mouchet, G.-K. Plattner, J. Sarmiento, R. Schlitzer, R. Slater, I. J. Totterdell, M.-F. Weirig, Y. Yamanaka and A. Yool (2002) Evaluation of ocean model ventilation with CFC-11: comparison of 13 global ocean models. *Ocean Modelling*, **4** (2), pp 89-120. doi: 10.1016/S1463-5003(01)00013-0
- Eyring, V., J. M. Arblaster, I. Cionni, J. Sedláček, J. Perlwitz, P. J. Young, S. Bekki, D. Bergmann, P. Cameron-Smith, W. J. Collins, G. Faluvegi, K.-D. Gottschaldt, L. W. Horowitz, D. E. Kinnison, J.-F. Lamarque, D. R. Marsh, D. Saint-Martin, D. T. Shindell, K. Sudo, S. Szopa and S. Watanabe (2013) Long-term ozone changes and associated climate impacts in CMIP5 simulations. *Journal of Geophysical Research: Atmospheres*, **118** (10), pp 5029-5060. doi: 10.1002/jgrd.50316
- Flato, G., J. Marotzke, B. Abiodun, P. Braconnot, S. C. Chou, W. Collins, P. Cox, F. Driouech, S. Emori, V. Eyring, C. Forest, P. Gleckler, E. Guilyardi, C. Jakob, V. Kattsov, C. Reason and M. Rummukainen and contributing authors (2013) Evaluation of Climate Models (Chapter 9). In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Stocker, T. F., D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P. M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp 741-866. doi: 10.1017/CBO9781107415324.020
- Forster, P. M., T. Andrews, P. Good, J. M. Gregory, L. S. Jackson and M. Zelinka (2013) Evaluating adjusted forcing and model spread for historical and future scenarios in the CMIP5 generation of climate models. *Journal of Geophysical Research*, **118** (3), pp 1139-1150. doi: 10.1002/jgrd.50174
- Gille, S. T. (2002) Warming of the Southern Ocean Since the 1950s. *Science*, **295** (5558), pp 1275-1277. doi: 10.1126/science.1065863
- Gille, S. T. (2008) Decadal-Scale Temperature Trends in the Southern Hemisphere Ocean. *Journal of Climate*, **21** (18), pp 4749-4765. doi: 10.1175/2008JCLI2131.1

725 Gleckler, P. J., B. D. Santer, C. M. Domingues, D. W. Pierce, T. P. Barnett, J. A. Church, K. E.
 726 Taylor, K. M. AchutaRao, T. P. Boyer, M. Ishii and P. M. Caldwell (2012) Human-induced global
 727 ocean warming on multidecadal timescales. *Nature Climate Change*, **2**, pp 524-529. doi:
 728 10.1038/NCLIMATE1553
 729
 730 Good, S. A. (2011) Depth Biases in XBT Data Diagnosed Using Bathymetry Data. *Journal of*
 731 *Atmospheric and Oceanic Technology*, **28** (2), pp 287-300. doi: 10.1175/2010JTECHO773.1
 732
 733 Good, S. A., M. J. Martin and N. A. Rayner (2013) EN4: Quality controlled ocean temperature and
 734 salinity profiles and monthly objective analyses with uncertainty estimates. *Journal of*
 735 *Geophysical Research: Oceans*, **118** (12), pp 6704-6716. doi: 10.1002/2013JC009067
 736
 737 Gouretski, V. and K. P. Koltermann (2007) How much is the ocean really warming? *Geophysical*
 738 *Research Letters*, **34**, L01610. doi: 10.1029/2006GL027834
 739
 740 Gouretski, V. and F. Reseghetti (2010) On depth and temperature biases in bathythermograph
 741 data: Development of a new correction scheme based on analysis of a global ocean database.
 742 *Deep Sea Research Part I*, **57** (6), pp 812-833. doi: 10.1016/j.dsr.2010.03.011
 743
 744 Gouretski, V. (2012) Using GEBCO digital bathymetry to infer depth biases in the XBT data. *Deep*
 745 *Sea Research Part I*, **62**, pp 40-52. doi: 10.1016/j.dsr.2011.12.012
 746
 747 Gouretski, V., J. Kennedy, T. Boyer and A. Köhl (2012) Consistent near-surface ocean warming
 748 since 1900 in two largely independent observing networks. *Geophysical Research Letters*, **39**
 749 (19), L19606. doi: 10.1029/2012GL052975
 750
 751 Gregory, J. M., H. T. Banks, P. A. Stott, J. A. Lowe and M. D. Palmer (2004) Simulated and
 752 observed decadal variability in ocean heat content. *Geophysical Research Letters*, **31**, L15312.
 753 doi: 10.1029/2004GL020258
 754
 755 Hamon, M., G. Reverdin and P.-Y. Le Traon (2012) Empirical Correction of XBT Data. *Journal of*
 756 *Atmospheric and Oceanic Technology*, **29** (7), pp 960-973. doi: 10.1175/JTECH-D-11-00129.1
 757
 758 Harrison, D. E. and M. Carson (2007) Is the World Ocean Warming? Upper-Ocean Temperature
 759 Trends: 1950-2000. *Journal of Physical Oceanography*, **37** (2), pp 174-187. doi:
 760 10.1175/JPO3005.1
 761
 762 Hu, A., and C. Deser (2013) Uncertainty in future regional sea level rise due to internal climate
 763 variability. *Geophysical Research Letters*, **40** (11), pp 2768-2772. doi: 10.1002/grl.50531
 764
 765 Ishii, M., M. Kimoto and M. Kachi (2003) Historical Ocean Subsurface Temperature Analysis with
 766 Error Estimates. *Monthly Weather Review*, **131** (1), pp 51-73. doi: 10.1175/1520-
 767 0493(2003)131<0051:HOSTAW>2.0.CO;2
 768
 769 Ishii, M., A. Shouji, S. Sugimoto and T. Matsumoto (2005) Objective analyses of sea-surface
 770 temperature and marine meteorological variables for the 20th century using ICOADS and the
 771 Kobe Collection. *International Journal of Climatology*, **25** (7), pp 865-879. doi: 10.1002/joc.1169

- Ishii, M., M. Kimoto, K. Sakamoto and S.-I. Iwasaki (2006) Steric Sea Level Changes Estimated from Historical Ocean Subsurface Temperature and Salinity Analyses. *Journal of Oceanography*, **62**, (2), pp 155-170. doi: 10.1007/s10872-006-0041-y
- Ishii, M. and M. Kimoto (2009) Reevaluation of Historical Ocean Heat Content Variations with Time-Varying XBT and MBT Depth Bias Corrections. *Journal of Oceanography*, **65** (3), pp 287-299. doi: 10.1007/s10872-009-0027-7
- Johnson, G. C., J. M. Lyman, J. K. Willis, S. Levitus, T. Boyer, J. Antonov, S. A. Good, C. M. Domingues, S. Wijffels and N. Bindoff (2013) Global oceans: Ocean heat content, in *State of the Climate in 2012*, **94**, [Blunden, J. and D. S. Arndt (eds.)]. Bulletin of the American Meteorological Society, pp 50-53. doi: 10.1175/2013BAMSStateoftheClimate.1
- Landerer, F. W., P. J. Gleckler and T. Lee (2013) Evaluation of CMIP5 dynamic sea surface height multi-model simulations against satellite observations. *Climate Dynamics*. **Early online release**. doi: 10.1007/s00382-013-1939-x
- Levitus, S. (1982) *Climatological Atlas of the World Ocean*. NOAA Professional Paper No. 13, U.S. Government Printing Office, Washington D.C., 173 pp
- Levitus, S. (1984) Annual Cycle of Temperature and Heat Storage in the World Ocean. *Journal of Physical Oceanography*, **14** (4), pp 727-746. doi: 10.1175/1520-0485(1984)014<0727:ACOTAH>2.0.CO;2
- Levitus, S., J. I. Antonov, T. P. Boyer and C. Stephens (2000) Warming of the World Ocean. *Science*, **287** (5461), pp 2225-2229. doi: 10.1126/science.287.5461.2225
- Levitus, S., J. Antonov and T. Boyer (2005) Warming of the world ocean, 1955-2003. *Geophysical Research Letters*, **32** (2), L02604. doi: 10.1029/2004GL021592
- Levitus, S., J. I. Antonov, T. P. Boyer, R. A. Locarnini, H. E. Garcia and A. V. Mishonov (2009) Global ocean heat content 1955-2008 in light of recently revealed instrumentation problems. *Geophysical Research Letters*, **36** (7), L07608. doi: 10.1029/2008GL037155
- Levitus, S., J. I. Antonov, T. P. Boyer, O. K. Baranova, H. E. Garcia, R. A. Locarnini, A. V. Mishonov, J. R. Reagan, D. Seidov, E. S. Yarosh and M. M. Zweng (2012) World ocean heat content and thermosteric sea level change (0-2000m), 1955-2010. *Geophysical Research Letters*, **39** (10), L10603. doi: 10.1029/2012GL051106
- Lyman, J. M. and G. C. Johnson (2008) Estimating Annual Global Upper-Ocean Heat Content Anomalies despite Irregular In Situ Ocean Sampling. *Journal of Climate*, **21** (21), pp 5629-5641. doi: 10.1175/2008JCLI2259.1
- Lyman, J. M., S. A. Good, V. V. Gouretski, M. Ishii, G. C. Johnson, M. D. Palmer, D. M. Smith and J. K. Willis (2010) Robust warming of the global upper ocean. *Nature*, **465** (7296), pp 334-337. doi: 10.1038/nature09043

817 Lyman, J. M. and G. C. Johnson (2014) Estimating Global Ocean Heat Content Changes in the
 818 Upper 1800 m since 1950 and the Influence of Climatology Choice. *Journal of Climate*, **27** (5), pp
 819 1945-1957. doi: 10.1175/JCLI-D-12-00752.1
 820
 821 Marcos, M. and A. Amores (2014) Quantifying anthropogenic and natural contributions to
 822 thermosteric sea level rise. *Geophysical Research Letters*, **41** (7), pp 2502-2507. doi:
 823 10.1002/2014GL059766
 824
 825 Meehl, G. A., G. J. Boer, C. Covey, M. Latif and R. J. Stouffer (2000) The Coupled Model
 826 Intercomparison Project (CMIP). *Bulletin of the American Meteorological Society*, **81** (2), pp 313-
 827 318. doi: 10.1175/1520-0477(2000)081<0313:TCMIPC>2.3.CO;2
 828
 829 Meehl, G. A., C. Covey, T. Delworth, M. Latif, B. McAvaney, J. F. B. Mitchell, R. J. Stouffer and K.
 830 E. Taylor (2007) The WCRP CMIP3 Multimodel Dataset. *Bulletin of the American Meteorological*
 831 *Society*, **88** (9), pp 1383-1394. doi: 10.1175/BAMS-88-9-1383
 832
 833 Meinshausen, M., S. J. Smith, K. Calvin, J. S. Daniel, M. L. T. Kainuma, J-F. Lamarque, K.
 834 Matsumoto, S. A. Montzka, S. C. B. Raper, K. Riahi, A. Thomson, G. J. M. Velders and D. P. P. van
 835 Vuuren (2011) The RCP greenhouse gas concentrations and their extensions from 1765 to 2300.
 836 *Climatic Change*, **109** (1), pp 213-241. doi: 10.1007/s10584-011-0156-z
 837
 838 Meyssignac, B., D. Salas y Melia, M. Becker, W. Llovel and A. Cazenave (2012) Tropical Pacific
 839 spatial trend patterns in observed sea level: internal variability and/or anthropogenic signature?
 840 *Climate of the Past*, **8**, pp 787-802. doi: 10.5194/cp-8-787-2012
 841
 842 Morrison, A. K., O. A. Saenko, A. McC. Hogg and P. Spence (2013) The role of vertical eddy flux in
 843 Southern Ocean heat uptake. *Geophysical Research Letters*, **40** (20), pp 5445-5450. doi:
 844 10.1002/2013GL057706
 845
 846 Nakićenović, N. and Swart, R. (Eds.) (2000) IPCC Special Report on Emissions Scenarios (SRES).
 847 Cambridge University Press, U.K., 570 pp. Available online:
 848 <https://www.ipcc.ch/ipccreports/sres/emission/>
 849
 850 Palmer, M. D., K. Haines, S. F. B. Tett and T. J. Ansell (2007) Isolating the signal of ocean global
 851 warming. *Geophysical Research Letters*, **34**, L23610. doi: 10.1029/2007GL031712
 852
 853 Palmer, M. D., S. A. Good, K. Haines, N. A. Rayner and P. A. Stott (2009) A new perspective on
 854 warming of the global oceans. *Geophysical Research Letters*, **36** (20), L20709. doi:
 855 10.1029/2009GL039491
 856
 857 Pierce, D. W., T. P. Barnett, K. M. AchutaRao, P. J. Gleckler, J. M. Gregory and W. M. Washington
 858 (2006) Anthropogenic Warming of the Ocean: Observations and Model Results. *Journal of*
 859 *Climate*, **19** (10), pp 1873-1900. doi: 10.1175/JCLI3723.1
 860
 861 Pierce, D. W., P. J. Gleckler, T. P. Barnett, B. D. Santer and P. J. Durack (2012) The fingerprint of
 862 human-induced changes in the ocean's salinity and temperature fields. *Geophysical Research*
 863 *Letters*, **39** (21). doi: 10.1029/2012GL053389

- Rahmstorf, S. (1995) Climate drift in an ocean model coupled to a simple, perfectly matched atmosphere. *Climate Dynamics*, **11** (8), pp 447-458. doi: 10.1007/BF00207194
- Rhein, M., S. R. Rintoul, S. Aoki, E. Campos, D. Chambers, R. A. Feely, S. Gulev, G. C. Johnson, S. A. Josey, A. Kostianoy, C. Mauritzen, D. Roemmich, L. D. Talley, F. Wang and contributing authors (2013) Observations: Ocean (Chapter 3). In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Stocker, T. F., D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P. M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp 255-315. doi: 10.1017/CBO9781107415324.010
- Sallée, J.-B., E. Shuckburgh, N. Bruneau, A. J. S. Meijers, T. J. Bracegirdle and Z. Wang (2013) Assessment of Southern Ocean mixed layer depths in CMIP5 models: Historical bias and forcing response. *Journal of Geophysical Research: Oceans*, **118** (4), pp 1845-1862. doi: 10.1002/jgrc.20157
- Schmidt, G. A., J. H. Jungclaus, C. M. Ammann, E. Bard, P. Braconnot, T. J. Crowley, G. Delaygue, F. Joos, N. A. Krivova, R. Muscheler, B. L. Otto-Bliesner, J. Pongratz, D. T. Shindell, S. K. Solanki, F. Steinhilber and L. E. A. Vieira (2012) Climate forcing reconstructions for use in PMIP simulations of the Last Millenium (v1.1). *Geoscientific Model Development*, **5**, pp 185-191. doi: 10.5194/gmd-5-185-2012
- Sen Gupta, A., L. C. Muir, J. N. Brown, S. J. Phipps, P. J. Durack, D. Monselesan and S. E. Wijffels (2012) Climate Drift in the CMIP3 Models. *Journal of Climate*, **25** (13), pp 4621-4640. doi: 10.1175/JCLI-D-11-00312.1
- Sen Gupta, A., N. C. Jourdain, J. N. Brown and D. Monselesan (2013) Climate Drift in the CMIP5 models. *Journal of Climate*, **26** (21), pp 8597-8615. doi: 10.1175/JCLI-D-12-00521.1
- Shao, A. E., S. Mecking, L. Thompson and R. E. Sonnerup (2013) Mixed layer saturations of CFC-11, CFC-12, and SF₆ in a global isopycnal model. *Journal of Geophysical Research*, **118** (10), pp 4978-4988. Doi: 10.1007/jgrc.20370
- Smith, D. M. and J. M. Murphy (2007) An objective ocean temperature and salinity analysis using covariances from a global climate model. *Journal of Geophysical Research*, **112** (C02022). doi: 10.1029/2005JC003172
- Taylor, K. E., R. J. Stouffer and G. A. Meehl (2012) An Overview of CMIP5 and the Experiment Design. *Bulletin of the American Meteorological Society*, **93** (4), pp 485-498. doi: 10.1175/BAMS-D-11-00094.1
- Vernier, J.-P., L. W. Thomason, J.-P. Pommereau, A. Bourassa, J. Pelon, A. Garnier, A. Hauchecorne, L. Blanot, C. Trepte, D. Degenstein and F. Vargas (2011) Major influence of tropical volcanic eruptions on the stratospheric aerosol layer during the last decade. *Geophysical Research Letters*, **38**, L12807. doi: 10.1029/2011GL047563
- von Schuckmann, K. and P.-Y. Le Traon (2011) How well can we derive Global Ocean Indicators from Argo data? *Ocean Science*, **7**, pp 783-791. doi: 10.5194/os-7-783-2011

912 Waugh, D. W., F. Primeau, T. DeVries and M. Holzer (2013) Recent Changes in the Ventilation of
 913 the Southern Oceans. *Science*, **568** (6119), pp 568-570. doi: 10.1126/science.1225411
 914
 915 Wijffels, S. E., J. Willis, C. M. Domingues, P. Barker, N. J. White, A. Gronell, K. Ridgway and J. A.
 916 Church (2008) Changing Expendable Bathythermograph Fall Rates and Their Impact on Estimates
 917 of Thermosteric Sea Level Rise. *Journal of Climate*, **21** (21), pp5657-5672. doi:
 918 10.1175/2008JCLI2290.1
 919
 920 Willis, J. K., D. Roemmich and B. Cornuelle (2004) Interannual variability in upper ocean heat
 921 content, temperature and thermosteric expansion on global scales. *Journal of Geophysical*
 922 *Research: Oceans*, **109** (C12). doi: 10.1029/2003JC002260
 923
 924 Yin, J. (2012) Century to multi-century sea level rise projections from CMIP5 models. *Geophysical*
 925 *Research Letters*, **39** (17), L17709. doi: 10.1029/2012GL052947
 926
 927 Zhang, X. and J. A. Church (2012) Sea level trends, Interannual and decadal variability in the
 928 Pacific Ocean. *Geophysical Research Letters*, **39** (21), L21701. doi: 10.1029/2012GL053240